THE

QUARTERLY JOURNAL
OF
EXPERIMENTAL PSYCHOLOGY

Volume VI

PUBLISHED FOR
THE EXPERIMENTAL PSYCHOLOGY GROUP
BY
W. HEFFER & SONS LTD. CAMBRIDGE

The Quarterly Journal of Experimental Psychology

VOLUME VI, 1954

CONTENTS

MISCELLANEA

BOOK REVIEWS

THE QUARTERLY JOURNAL OF
EXPERIMENTAL PSYCHOLOGY

Vol. VI 1954 Part 1

SOME EFFECTS OF NOISE ON VISUAL PERFORMANCE

BY

D. E. BROADBENT

From the Medical Research Council Applied Psychology Research Unit, Cambridge.

A group of ten subjects showed impaired performance, when watch-keeping on a display made up of steam-pressure gauges, in 100 db. noise as compared with 70 db. On the easier task of watch-keeping on a display made up of small lights, another group of twenty subjects showed no overall effect of noise. Individual subjects who showed a practice effect on the latter task comparable to that shown by all subjects on the former one, however, also showed a similar effect of noise.

In addition, performance on the light-watching became relatively less efficient in noise with continued exposure: and although parts of the task were still adequately carried out, others were not. The fact that noise effects are thus functions of individual differences, of visibility of signal, and of length of performance in noise, allows us to explain the negative findings of many previous workers.

I

Introduction

Recent reviews of the work on the psychological effects of noise (Berrien, 1946; Kryter, 1950, 1952) state correctly that there is no adequate existing evidence of any effect beyond the initial disturbance produced by startling noises. Yet as Konorski (1948) points out, extremely valuable information about the nature of processes in the nervous system is likely to be derived from interference and facilitation between them. If those who complain of effects of noise have objective grounds for doing so, it may be that theories of behaviour should consider the rôle of background stimulation more thoroughly than they do. The factual issue is thus of some importance, and the following experiments were designed to cover certain possibilities which the existing work still left open. These possibilities may be summarized thus:

1. The effect of task difficulty. It has been shown by Mackworth (1950, p. 57) that "fatigue" effects are more easily shown with faint signals than with easily visible ones, and a similar relation may apply to other forms of stress.

2. The rôle of individual differences. It has been found by the writer in preliminary experiments that "fatigue" effects appear more readily in certain individuals, who may be detected by the size of the practice effect between runs. A similar relation may apply to other forms of stress.

3. The effect of prolonged performance in noise. It is in general true (Babington-Smith, 1951) that frequent changes in conditions do not allow human performance to settle to a level characteristic of each condition: prolonged performance in noise may therefore be needed to show decrement.

4. The use of analytic rather than overall scores. Complex tasks may give similar overall scores under stress although the relation between parts of the performance may in fact be changing: a point put for instance by Bartlett (1945).

II

Procedure

Production of noise. A room was set up sixteen feet square and eight feet high, built of metal plate loaded to prevent resonance. In three of the walls loudspeakers were inserted, each wall carrying one H/F speaker with a multi-cell horn and one L/F speaker with re-entrant horn. The frequency of cross-over was 1000 cycles. The speakers were fed from an R.G.D. tape recorder through 50-watt Vortexion amplifiers, one for each pair of speakers. This arrangement gave a flat field over the whole area in which the subject worked.

The noise itself was recorded on an endless loop of tape, the recording having been made near rapidly moving machinery. The spectrum was reasonably flat ($\pm$ 3 db.) on third-octave analysis from 40–5000 c.p.s. On "noise" days the level used was 100 db. re 0·0002 dynes/sq. cm., and on "quiet" days the same recording was played at 70 db. to mask minor extraneous noises. The levels were checked daily by a General Radio sound-level meter which was itself checked at intervals against Admiralty Research Laboratory equipment.

The tasks employed. Two visual vigilance tasks were given to separate groups of subjects. The first, the Twenty Dials Test, required the subject to watch twenty steam-pressure gauges: and if he saw any of the pointers reading above a danger mark, to turn a knob below the gauge to bring the pointer back below the mark. The gauges were arranged on three sides of a square of 12 ft. side, the subject being seated in the middle of the open side. Each gauge was six inches in diameter; the pointers were stationary except when a signal was delivered, when the tip of the pointer moved through 0·8 in to a new position above the danger mark and again became stationary. A silent mechanism operating on the principle of the ammeter was employed to move the pointer. The length of watch was one and a half hours, and fifteen signals were delivered in this time at intervals varying from 1 to 12 minutes. Each half-hour of the watch contained a signal at each of the intervals used, and position of the signal in the display was similarly counterbalanced.

In various preliminary experiments on sixty subjects, it was found that the first watch on this task was anomalous due to practice, but that when it is repeated five times at twenty-four hour intervals under quiet conditions the last four watches do not vary significantly in performance. If anything, performance is better on the third and fourth days. Consequently for this investigation ten subjects were given the task five times, with "noise" on the third and fourth days and the remaining days "quiet."

The second task, the Twenty Lights Test, was similar to the Twenty Dials except that each dial was replaced by a head-lamp bulb run at half-voltage and covered with paper so as to give a dim glow. Each bulb could be clearly recognised as "on" or "off" in foveal vision, but did not attract attention in the periphery: their brightness was approximately 5 foot-lamberts against a background of 3 foot-lamberts. To prevent the task becoming too easy, the Lights were not all at eye-level as the Dials were, but irregularly arranged from ground-level to a height of four feet. The response in this case was to press a key when a lamp was seen alight, the distance of the key from the subject's chair being the same as that of the knobs in the Dials Test. The other details of administration were the same as for the Dials.

Ten subjects were given this task five times each, with noise on the third and fourth days: since this task had not been used repeatedly in quiet like the Dials, a further ten subjects were given the task five times, with noise on the first and second days.

Both on the Dials and Lights, the experimenter and controlling apparatus were in a control room separated from the metal chamber by an air space and a brick wall. As a check on wakefulness, etc., the subject could be observed from time to time without his knowledge. All subjects were Naval Ratings aged 18–30 years. It is important to note that almost all of them were familiar with noise levels at least as high as the one used from their ordinary work, but that this did not prevent effects from appearing. Twenty of them were Fleet Air Arm ratings (mostly mechanics), three were stokers, four seamen, and three radar operators.

III

Results

Effect of task difficulty. Responses on the Dials are divided into "seens" and "founds," the former being those in which the subject reports that he was looking at the dial at the instant when the pointer moved, and so had his attention caught by

this conspicuous signal. These show no significant effect of noise. In the "founds," the proportion responded to in nine seconds or less is taken as the "quick found" score. The mean proportion of "quick founds" for each run is shown in Table I. As in the earlier work, there is a practice effect between runs 1 and 2. The difference between runs 3 and 4 in noise and runs 2 and 5 in quiet is significant, $t = 3 \cdot 21$ while the 0·05 level is 2·26 and the 0·01 level is 3·25.

On the Lights, the proportion of "quick founds" is higher throughout, the signal being a much more conspicuous one. On the group given the same testing schedule as the Dials group, there is no practice effect between runs 1 and 2. Nor is there any difference between noise and quiet in either of the two Lights groups. The more difficult signal, therefore, shows an effect of noise while the easier one does not.

TABLE I

THE PROPORTION OF "QUICK FOUNDS" IN NOISE AND QUIET

Task and Group	Run				
	1 *Quiet (Practice)*	2 *Quiet*	3 *Noise*	4 *Noise*	5 *Quiet*
Dials	0·215	0·339	0·193	0·244	0·381
Lights: the markedly improving subjects ..	0·424	0·574	0·328	0·374	0·454
Lights: the least improving subjects..	0·432	0·226	0·482	0·300	0·414

NOTE:—On the Lights Test the subjects shown are those treated similarly to the Dials subjects, so as to facilitate comparison. See text for the control of order effects by other groups. The "markedly improving" subjects are the five showing the greatest value of (Run 2—Run 1) and the "least improving" the five remaining.

Effect of individual differences. Despite the absence of any overall effect on the Lights, individual differences appear which can be best illustrated comparatively by the method of Table I. Here the ten subjects treated similarly to the Dials group are split into the five showing the greatest and the five showing the least improvement from run 1 to 2. The effect for the former is strikingly parallel to that for the Dials, while the latter behave quite differently. To test the significance of this difference, we may consider all twenty subjects tested on the Lights, neglecting the last quiet run (which was given to ensure that all subjects had similar expectancies). Each subject then has two successive runs under each condition, half the subjects having noise first and half quiet. If we then correlate improvement between the two quiet runs with the amount of deterioration in noise, i.e.

$$(Q_2 - Q_1) \text{ and } (Q_2 + Q_1) - (N_2 + N_1)$$

we obtain a tau summed for the two groups of 0·36, which is significant, S/S.E. of $S = 2 \cdot 09$, while the 0·05 level is 1·96. This correlation is not due to exaggerated random variation in Q_2: both $(Q_2 - N_2)$ and $(Q_1 - N_1)$ correlate to approximately the same extent with $(Q_2 - Q_1)$. An independent test has also been used to differentiate the susceptible individuals: see Discussion.

Effect of prolonged performance. In Table I we see the difference between quiet and noise performance on the Lights for each successive half-hour from the beginning of a particular condition. It will be noted that a slight initial decrement in noise is followed by an actual advantage for noise, which decreases progressively during the remaining half-hours. It is not possible to obtain an adequate estimate of error for each of these differences, since this experimental design (like any other) must confound certain interactions involving subjects which are known from the previous work to be important. By taking the scores for whole runs rather than half-hours, however, most of these are eliminated, and it then appears that $N_1 - N_2$ is very significantly greater than $Q_1 - Q_2$; that is, the general downward trend of Table I is confirmed. $F = 8{\cdot}46$ while the 0·01 level is 8·28. Performance in noise becomes worse relative to that in quiet with increasing length of exposure. It may be noted that this fact is consistent with the interpretation that the apparent absence of overall effect on the Lights is due to a compensatory effort by some subjects, and that this effort cannot be maintained.

TABLE II

DIFFERENCE BETWEEN NOISE AND QUIET ON THE LIGHTS

Half-hours since beginning of condition	1	2	3	Gap between runs	4	5	6
Noise-quiet (proportion of quick founds)	−0·046	+0·140	+0·058	—	−0·004	−0·101	−0·133

Use of analytic scores. This downward movement does not reach a point at which performance on the Lights is significantly worse in noise than in quiet, using the score so far employed. But certain component scores do show a change in the pattern of performance in noise. Thus, if we take the longest succession of signals each of which took more than 9 seconds to elicit a response, we find that the series is longer in the second noise run than in the second quiet run. Using a logarithmic transformation, $F = 5{\cdot}18$ while the 0·05 level is 4·41. The length of the sequence of misses corresponding to the mean of the transformed series is three signals for quiet and four for noise. Missed signals appear therefore to be more concentrated by noise into certain periods of the watch.

There is also a tendency for the lights in front of the subject to be neglected as compared with those at the sides. Considering the number of signals from the lights facing the subject responded to in 4 seconds or less, there is a mean advantage for the second quiet run as opposed to the second noise run of 0·6 signals, for which $t = 2{\cdot}45$ while the 0·05 level is 2·09. Similar responses to the lights on each side of the subject showed a slight improvement in noise, but quite insignificant ($t = 1{\cdot}28$)

Thus it appears that parts of the performance both in space and time are impaired by noise while other parts are not; so that overall scores show no significant effect, being diluted by the unaffected parts.

IV

Discussion

In certain respects (task difficulty and the use of analytic scores) the present results show reasons which might explain the negative findings of the most careful previous workers: but these factors may or may not have been important in earlier

experiments. Other points, however, were quite certainly operative. Most important of these are the individual differences.

The individuals most susceptible to noise show the largest practice effects between the first two runs in quiet: and it may be noted that in the preliminary experiments it was found that this score also identified the subjects showing the largest "fatigue" effects. An independent test was found to correlate with fatigability both on the Dials and Lights, and also to show a relation to the effect of noise. This test was unrelated to test intelligence but showed significant links with academic attainment. It was a modification of the Level of Aspiration test used by Eysenck and his collaborators (1947) to distinguish dysthymics and hysterics, a dimension which is known to be related to education. It seems likely, then, that a sample of subjects selected by academic examination or similar means will also by that very fact be selected for resistance to the effects of continuous performance and of irrelevant attention-demanding stimulation. The present research was more fortunate than almost all its predecessors in having available a sample not drawn from the undergraduate or research worker population.

Secondly, the length of performance used in these experiments differs from earlier work. On the one hand tasks as long or longer have been performed with occasional short bursts of noise, and on the other hand very long exposures to noise have been investigated by short tests. But it is unusual to find previous work of comparable scope on the performance of one task in one condition of noise or quiet for periods as long as those used here. Yet, as the data of Table II show, this is clearly likely to falsify results. When a prolonged task and a long exposure were used by Davis (1948), the results suggested a real but complex effect of noise with marked individual differences.

Finally, and perhaps most important of these factors, is the nature of vigilance tasks themselves. It is likely that tasks requiring reception of information from a particular source at an uncertain time are peculiarly vulnerable to the effects of intense competing stimuli such as noise. This peculiarity of vigilance tasks requires further experimental evidence for its elucidation, and has been examined elsewhere (Broadbent, 1953). We may note here simply that earlier work on noise has not used this type of task.

This work was supported financially by the Medical Research Council, and thanks are due to the Royal Navy both for supplying subjects and for providing the very extensive technical facilities required. The influence of Professor Sir Frederic Bartlett, F.R.S., and of Dr. N. H. Mackworth will be readily apparent throughout the paper.

References

Babington-Smith, B. (1951). On some difficulties encountered in the use of factorial designs and analysis of variance in psychological experiments. *Brit. J. Psychol.*, **42**, 250–268.

Bartlett, F. C. (1945). Some growing points in experimental psychology. *Endeavour*, **4**, No. 14.

Berrien, F. K. (1946). The effects of noise. *Psychol. Bull.*, **43**, 141–161.

Broadbent, D. E. (1953). Noise, paced performance and vigilance tasks. *Brit. J. Psychol.*, **44**, 295–303.

Davis, D. R. (1948). Pilot Error. A.P. 3139a. London: H.M. Stationery Office.

Eysenck, H. J. (1947). *Dimensions of personality*. London.

Konorski, J. (1948). *Conditioned reflexes and neuron organisation*. Cambridge.

Kryter, K. D. (1950). The effects of noise on man. *J. Speech Dis., Monogr. Suppl.* No. 1.

Kryter, K. D. (1952). Noise safety criteria. *Arch. industr. Hyg.*, **5**, 117–120.

Mackworth, N. H. (1950). *Researches on the Measurement of Human Performance.* Medical Research Council Special Report No. 268. London: H.M. Stationery Office.

A MACHINE WITH INSIGHT

BY

J. A. DEUTSCH

From the Institute of Experimental Psychology, Oxford

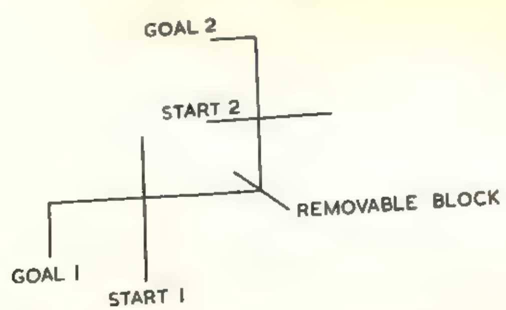
When an animal solves a novel problem without trial and error, a psychologist tends to call the behaviour insightful. A relatively simple machine resting on new principles, capable of this and learning, has been constructed by the writer in order to demonstrate his tentative explanation of this and other aspects of animal behaviour.

The machine consists of three parts. The first retains information and utilises it in accordance with the goal set. The second is a trolley which is guided by a pulse which the first part transmits; it also has bumpers which cause it reflexly to steer away from any obstacle and to turn out of corners by reversing one of the motors which drive each of the two main wheels separately; and it has a light mounted on it which is thrown forward at a wide angle. The third part comprises photocells, which act as receptors and are attached to the walls of the maze which the trolley is required to learn.

Behaviour

The machine's performance is more precisely described as follows. It learns and retains any two short, modified, rat mazes. On the learning run the machine is made to enter all the blind alleys of the maze on its way to the goal box. On the next run it will travel straight from the entrance of the maze to the goal box without entering a blind alley. If the two mazes, which it has learnt and retained, share a common point (even if this is down blind alleys in each maze), the machine will, on being set in an entrance, find its way without any further trial and error to whichever of the two goal boxes the demonstrator makes it seek. The machine appears to have this property in common with the rat, which suddenly integrates its past experience in accordance with its aim in the latent learning situation when a goal object is introduced.

FIGURE 1



After learning two mazes separately, the machine will solve the problem of finding its way from the entrance of one maze to the goal of the other.

Furthermore, the machine can transfer or generalize its "knowledge" to mazes of completely different shape and similar only in a highly abstract way. Each alley in the maze is marked by a different signal which is supplied to the machine. If the sequential order of the signals down the alleys with respect to each other is kept the same, even though the shape of the maze is quite altered, the machine can still find its way. It will execute an errorless run even though it may now have to

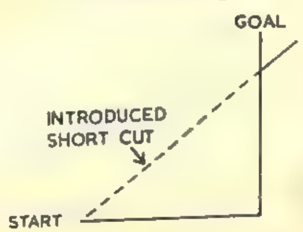
turn in opposite directions at the choice points. The machine does not learn a sequence of responses to a chain of appropriate stimuli. The movements that it makes are of little consequence both in its learning and its performance are always highly variable. It can be "taught" though the trolley be motionless and the machine be "paralysed" and make no motor response. The trolley will make an errorless run even when it is partially crippled.

FIGURE 2

The machine transfers its training from A to B and vice versa, but not from A to C.

The flexibility and variability of its performance may be illustrated by its ability to take advantage of short cuts.

FIGURE 3



The machine will take a short cut if it is made available.

After learning a maze, the machine will take a short cut, if one is introduced leading to a point nearer the goal. Similarly, it will take short cuts introduced when it is finding its way from the entrance of one maze to the goal in a second.

The machine can also be trained in operant conditioning situations, multiple choice discriminations and related problems.

The system

The behaviour described above is the behaviour of a certain type of system. The properties described are not those of relays but of arrangements of relays. Further, we can make the same arrangement, manifesting the same behaviour, out of other components—thermionic valves, transistors or wheels. Perhaps a similar arrangement is to be found in the mammalian central nervous system. After all, the mammal manifests similar behaviour.

The abstract system itself which is richer is described elsewhere (Deutsch, 1953). Here are described the operations which the machine performs and its particular embodiment as represented by this machine.

On the learning trial (before the goal is found) the system arranges the receptors in an order inverse to that in which they were stimulated. The receptor which was set off first will be last in the series, and the last one next to the first. The first position is always occupied by the goal-signalling receptor. There will therefore be a record of the order in which certain aspects of the environment occur in relation to the goal.

This order is preserved by making a pulse from a receptor travel to the trolley through a contact on relay with what might be called a particular number. The relay with the lowest number has the goal-signalling receptor already attached to it, and when it is switched on and closed by the operator, all the relays with higher numbers are closed, enabling any receptor attached to any of them to control the trolley, whichever it is that may be stimulated.

But when a pulse passes through any relay, all those relays which have a higher number will open and stay open, so that the receptors attached to them will not be able to influence the movements of the trolley. This will be called the hierarchical arrangement. It ensures that the trolley will always steer towards that "landmark" which it "saw" nearest (in time) to the goal.

When there are two series of receptors made, each registering the order of "cues" leading to the two goals, if there is a receptor common to both series, the system will exploit this information. When one of the goal-signal receptors is closed by the operator, it will be recalled that all the relays with higher numbers will also be switched on. Now one of the relays with a higher number has a receptor common to both series attached to it; through this common receptor will now flow excitation closing all the relays in the other series with a higher number. This gives the system "insight."

The machine

The system is embodied in the following way. First is the "thinking" part, composed of six uniselectors and twenty-three relays. The second is the trolley which moves through the maze, and the third consists of the photocell-receptors. The trolley has a light mounted on it, which is thrown forward at a wide angle, and which touches off the photocells mounted on each sector of the maze. (This arrangement was adopted as a practically easier equivalent of mounting receptors on the trolley and placing the signals in the maze.) The trolley has two responses: it either goes forward or turns round on its axis. When it has been set to learn its way to a goal, it goes forward (except when it bumps into the walls of the maze). When it has been set to find its way after it has learnt, it will continue to travel forward as long as a particular receptor which the machine has selected is being stimulated. When this stimulation ceases, the machine will turn until the stimulation of this receptor occurs again. It will thus guide itself towards the selected receptor. This is the way in which the trolley is controlled.

The way in which the machine stores the order in which the receptors were stimulated on the learning trial is now described. (See Figure 4.)

Each receptor is connected to its uniselector before it has been stimulated, in such a way that its stimulation will move the wipers of its own uniselector to the first position and the wipers of any other uniselector whose receptor has already been stimulated to the next position. Once it has been stimulated and the wipers shifted to the first position, the stimulation from its own receptor will not affect it; nor will it affect any other uniselector. It will now only be shifted by a receptor which is being stimulated for the first time.

This will be made clear if we study Figure 4. The uniselector U1, which is moved when receptor 1 is stimulated, has two principal states.

1. The state before receptor 1 has been stimulated. Here only the closing of relay R1, caused by the stimulation of receptor 1, will move the wipers of U1 from position O. Any pulse from R2, closed by the stimulation of receptor 2, will not affect U1 because the circuit which effects this is broken at U1, row b.

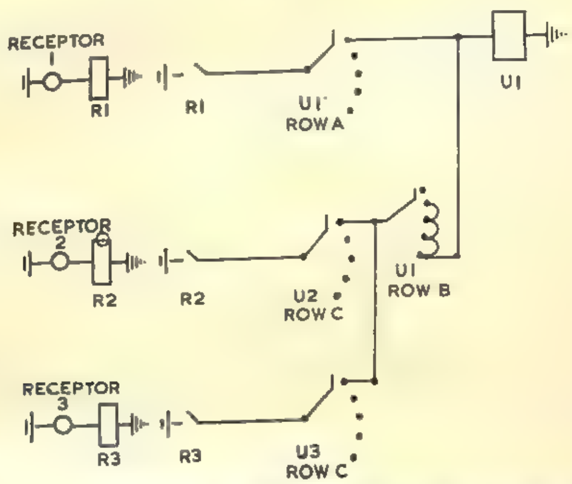
2. The state after receptor 1 has been stimulated. Here the closing of R1 will not further affect U1, because the wipers have now been shifted to position 1. And thus the circuit which shifts U1 when R1 closes will have been broken.

On the other hand the circuit which connects U1 with R2 and R3 is now closed by U1, row b. U1 will therefore be stepped on when R2 or R3 close. But U1 will only be affected if R2 (or R3) close for the first time because the circuit leading through U2 row c (or U3, row c) will be broken as soon as the wipers have been moved from position O.

When considering receptor n, substitute n for 1 in the above description and vice versa.

Now, if there are three receptors stimulated in turn, it follows that the uniselector attached to the first one to be stimulated will be on the third position, the next on the second position and the last on the first position.

FIGURE 4



The circuit used for attaching receptor 1 in its place during learning.

For receptor 2, exactly the same circuit is used, but 2 should be read in place of 1, and vice versa.

The receptors are also connected through wipers to the positions, each of which is separately attached to the hierarchical arrangement. Thus the nearer in time to the goal a receptor was touched off, the higher in the hierarchy is it connected. Once the goal the machine was set is found, the connection becomes permanent. As on the learning run the trolley is made to "explore" each alley or sector in turn before it reaches the goal, the receptors are connected to the hierarchical arrangement in accordance with their proximity to the goal-receptor.

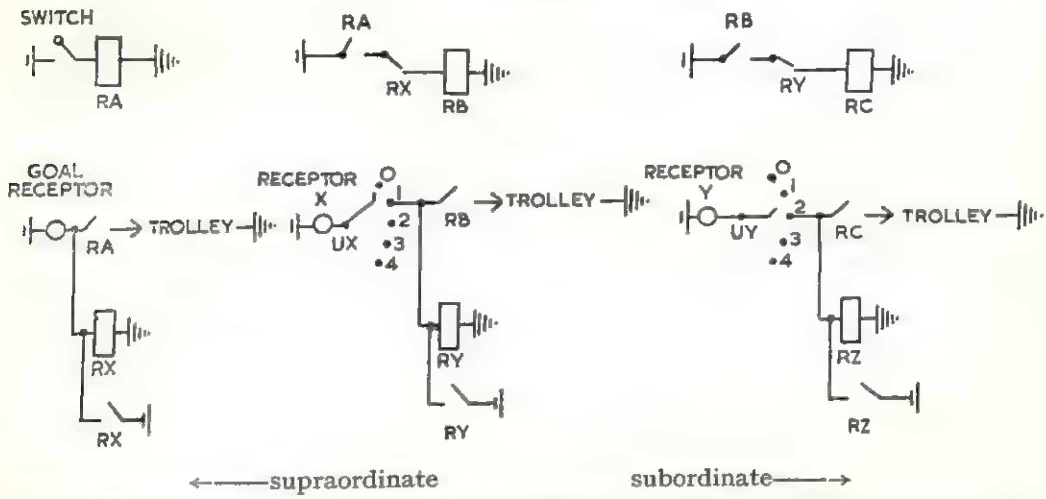
The order thus preserved is utilised in the following manner. (See Figure 5).

The receptors are connected by uniselectors to a hierarchical arrangement of relays. By this is meant that the relays are connected in a series in which the energizing of a preceding (supraordinate) relay causes the closing of its successor (subordinate), and the switching off a predecessor leads to the opening of the relay whose closing it previously caused.

The receptor whose stimulation functions as goal (two in the present model) is "innately" connected to the "make" contacts of the first relay in this series (RA) (again two in the present model) (see Figure 5). The rest of the receptors (x, y) are also connected to the "make" contacts of relays but each is connected to its respective relay in the hierarchy through the activity of its uniselector (Ux, Uy) during the "training" trial. Each receptor is connected at the same time to close another relay (RX, RY, RZ) which breaks the circuit of the relay subordinate to the one to which the receptor is also connected by the uniselector. When the hierarchy is switched on ("when there is a drive acting on the organism") the stimulation of any receptor connected to the hierarchy will cause the trolley to go forward while this stimulation continues, because it is connected to the trolley through the "make" contact on its relay in the hierarchy and secondly, will immediately lead to the opening, for the rest of the trial, of all the relays subordinate to that to which it is connected.

During the "training" trial the hierarchy is not switched on. Instead the circuit which allows the receptors to drive the uniselectors is closed, to be opened when the finding of the goal receptor is signalled.

FIGURE 5



The so-called hierarchical arrangement which selects the "cue" which the trolley is to approach. The make contact of relay RB is connected to the first position of every uniselector, (the previous position is designated O, because it leaves a receptor unconnected). RC is connected to the second, and so on.

On being set in the entrance of the maze on the test trial the machine is set to find the goal. The hierarchical arrangement of relays is now switched on and some receptors again pick up signals.

As has been explained, the receptors when they are stimulated send a pulse to the trolley through that portion of the hierarchy to which they are attached, but only when this portion is switched on at the same time. A portion is switched off as soon as another portion supraordinate to it has passed a pulse from the receptor, attached to it, to the trolley. Hence the trolley will always steer towards the receptor nearest to the goal and bring itself by this means within the range of stimulation of others even nearer until it reaches the goal receptor.

When the trolley has learnt two mazes with a common receptor and two separate goals the one receptor is attached to both hierarchies perhaps with a different position in each. When either goal is then selected, the switching on of the hierarchy

subordinate to that goal will be transmitted through the common receptor to portions of the other hierarchy. Only those portions, subordinate to that to which this common receptor is connected, will be switched on. Hence the trolley will find the shortest way to the goal selected in whichever entrance it is set.

The system consists of six uniselectors and twenty-three relays. The uniselectors have six wipers each, and are ordinary P.O. equipment. Only three positions are employed on each. Smaller ones would thus be more suitable, were they obtainable. The relays are similarly large P.O. relays with four "make" and one "break" contact. These also are often not all used, but the writer and constructor was led to use them because they were cheap and readily available and because he believes that he would have become intolerably confused had he used more than one type.

The system is such that it could be expanded without a loss in efficiency. For instance, even the time taken to calculate a choice or reach a decision (when there is no conflict) does not vary with the complexity of the operation (and the size of the system); it takes the machine no longer to select the path to a goal whether it has to combine information acquired in separate situations or not. This, with the number of complex properties which this relatively simple system has in common with a mammalian organism is encouraging from the point of view of theoretical psychology.

Reference

Deutsch, J. A. (1953). A new type of behaviour theory. *Brit. J. Psychol.*, **44**, 304–317.

FRUSTRATION AND STEREOTYPED BEHAVIOUR IN HUMAN SUBJECTS

BY

LEONARD C. T. JONES

From the Psychological Department, University College, London.

The investigation was designed to show the effects upon behaviour of three different durations of frustration, and two degrees of motivation during the frustrating period. Frustration was induced in 144 subjects by setting them the task of "learning" an insoluble temporal maze; they had to record their responses by pressing on one or other of two morse-type keys. Its effect was measured in terms of:

(1) The time taken to learn a soluble maze introduced, without the subjects' knowing it, by changing the system of "rewarding" responses from one based on chance to one based on the constant repetition of a pattern requiring the responses: Left—Right—Right, for its solution.

(2) The tendency for the subjects to show stereotypy of behaviour by responding on the same key for a number of trials in succession without reaching a solution.

(3) The pressure exerted on the response keys, which was taken to be a measure of vacillation.

Predictions that the time taken to learn the soluble task, and the stereotypy, would increase in direct relation to the duration of the frustrating period and the degree of motivation were tested.

It was found that, while there was an immediate increase in both the time taken to learn the soluble task and the stereotypy after a short period of frustration, a point was reached under conditions of prolonged frustration after which no further increase occurred and some adjustment to the situation was shown.

Some confirmation of the effect of increase in motivation in the direction predicted was obtained in all cases, but the differences were not statistically significant. The rankings of the subjects according to the time taken to learn the soluble task and the degree of stereotypy were found to be closely correlated. A definite tendency towards increased vacillation of response was seen in many of the records during the period when frustration might have been expected to have been at its peak.

These findings are discussed in relation to Maier's theory of frustration and to Selye's concept of a "general adaptation syndrome." The latter theory is more suitable for the interpretation of the results of the present investigation.

I

Introduction

The study of behaviour under conditions of stress and frustration has assumed considerable importance in behaviour theory, especially in relation to the problem of explaining behavioural maladaptation. In the present investigation frustration was produced in a situation in which response possibilities were limited and easily measurable. Frustration was defined as an intervening variable, the definition being taken from Underwood (1949):

> "Frustration is that which leads to the deviant behaviour which is observed as a result of the blocking of, or interference with, a goal-directed behaviour sequence."

With such a definition, an individual can only be said to be frustrated when the blocking of a drive by some obstacle leads to some form of "deviant" behaviour, or to a change in the anxiety level of the individual.

Previous investigations have shown that the behaviour resulting from frustration shows two main characteristics; firstly, it is more *stereotyped*, showing less plasticity

in a learning situation, and secondly, this stereotypy tends to persist during the period of frustration and for some time afterwards, that is to say, it shows signs of *rigidity*. But these are also characteristics of behaviour occurring as a result of any form of stress, and for this reason no stress-producing stimuli, other than the frustrating situation were employed, and in particular, the subjects' performance was in no way criticized, disparaged or ridiculed.

The insoluble problem was used as a means of producing frustration, and the performances of groups of subjects at a learning task following the frustrating situation were compared for three different durations of the insoluble problem and two degrees of motivation of the subjects during the frustrating period. Behaviour deviations were shown to occur under the conditions of stress produced by the problem, but under conditions of prolonged stress a diminished degree of stereotypy suggests that some adjustment was made to the situation.

II

Research Design and Procedure

Research Design.

The research design is presented in Table I and consisted of the following phases:

(1) A preliminary trial period of thirty seconds during which the problem was insoluble (five seconds were allowed for each response, so that six responses could be made in this first phase). Responses were "rewarded" on a random basis. The first phase gave the subjects the opportunity to become familiar with the apparatus and experimental situation, but it was essential that it should be short lest it should itself prove to be frustrating.

TABLE I

Research Design

Group	N	Phase 1	Phase 2	Phase 3		Phase 4
I	24	Trial period	Rest period	Insoluble problem: Non-incentive	5 trials	Soluble problem
II	24	"		Insoluble problem: Incentive	5 trials	"
III	24	"		Insoluble problem: Non-incentive	50 trials	"
IV	24	"		Insoluble problem: Incentive	50 trials	"
V	24	"		Insoluble problem: Non-incentive	100 trials	"
VI	24	"		Insoluble problem: Incentive	100 trials	"

Three different effects of frustration on behaviour were analysed: (i) the time taken to learn the maze following the frustrating situation (Phase 4); (ii) the degree of stereotypy of behaviour throughout phases 3 and 4 of the experiment. A succession of six responses or more to the same key was considered to be evidence of stereotypy; this criterion was chosen because under conditions of normal learning without any prior frustrating situation it was found that nearly all subjects could solve the problem after responding on one key for this number of trials; (iii) the degree of vacillation in making responses during phases 3 and 4, measured by variations in the pressure of the subjects' hand on the response keys.

(2) A rest period of two minutes during which the subjects were allowed to ask any questions about the procedure which had not been explained sufficiently in the instructions.

(3) An insoluble problem phase in which subjects of three main groups were given an insoluble temporal maze of durations 5, 50 and 100 trials, respectively, responses being "rewarded" on a random basis. Each group of subjects was sub-divided, members of each sub-group being given different instructions designed to create differing degrees of motivation.

(4) A soluble problem phase, the soluble problem being introduced, without the subject's knowing it, by changing the method of rewarding responses from a randomly determined sequence to one based on the constant repetition of a pattern requiring the responses: Left—Right—Right, for its solution. The criterion for the learning of the problem was five correct repetitions of the response pattern: Left—Right—Right; i.e. fifteen responses in all.

The following hypotheses were tested:—

(1) The length of time which it takes to solve a problem following a period of frustration increases in direct relation to the duration of the period of frustration.

(2) The length of time which it takes to solve a problem following a period of frustration increases in direct relation to the degree of motivation of the subject during the period of frustration.

(3) The degree of stereotypy of behaviour increases in direct relation to the duration of the period of frustration.

(4) The degree of stereotypy of behaviour increases in direct relation to the degree of motivation of the subject during the period of frustration.

Apparatus.

The subject was seated before a screen on which a red and a white light were displayed; fixed to a table in front of the screen was the apparatus on which his responses were made. The circuit for the red light was completed and broken every five seconds by a timer, this light acting as a signal for the subject to make a response. If a correct response was made—in accordance with the pattern which was set by the experimenter—the white light flashed on; if an incorrect response, a buzzer sounded. On the assumption that the light and buzzer gave information as to the pattern of responses to be learned, the subject attempted to eliminate those responses which resulted in the buzzer.

Responses were made by pressing down one or other of two keys which were built into a box along the top of which the subject could rest his arm comfortably, with the first and fourth fingers (forefinger and little-finger) on the two keys which protruded from the end of the apparatus. The keys were supported on a movable platform, concealed in the apparatus, which rested on a rubber bulb. Variations in the pressure exerted on the keys were conveyed through this bulb by means of a rubber tube to a tambour and pen recording on a drum-kymograph; the maximum distance through which the response keys could move was approximately two inches. The subjects rested their fingers on the keys throughout the experiment and so a continuous record of the degree of vacillation during each response was kept.

The two response keys were connected to the light and buzzer, respectively, through a mercury reversing switch by means of which the problem to be learned was set. A full record of the responses made—i.e. left or right—for each trial was kept by two pens of a four-pen tape polygraph. The third pen recorded the number of responses made by the subject for which the buzzer sounded, those responses indicating errors to the subject, while the fourth pen recorded the interval between each flash of the red light.

Subjects.

One hundred and forty-four subjects were assigned on a random basis to the six experimental groups each of which contained an equal number of men and women. The subjects were undergraduates of the University of London who volunteered to co-operate in the research and for this reason it may be assumed that they represent a fairly homogeneous group so far as intelligence level and socio-economic status were concerned. Very few of the subjects were taking courses in psychology or the social sciences. The median age of subjects was 21 years, with a range of 17 to 32 years.

Procedure.

At the outset of the experimental session one of the following two sets of instructions was given verbally to the subject, according to whether he had been assigned to a "non-incentive" or an "incentive" group.

Non-incentive groups:

"I am going to set you a problem on the apparatus in front of you. This problem will be used on other people and at present I am only concerned with finding its difficulty by trying it out on a small sample group.

"I want you to rest your arm on the box in front of you with your first and fourth fingers resting on the two keys (demonstrated by the experimenter). Your fingers should be resting on the keys the whole time. Every five seconds the red light on the screen will flash on. Each time this happens I want you to press down one or other of the two keys. Your job will be to learn a pattern which repeats itself continuously and which I am going to set you on those two keys. If you press down the correct key in accordance with this pattern, the white light in the centre of the screen will come on, if you press the wrong key then a buzzer will sound. Try to learn the pattern as quickly as possible, there is no time limit."

Incentive groups:

"This is an intelligence test which is often used in selecting college students in the United States.

"I want you to rest your arm on the box in front of you . . . (The instructions continue as above, but conclude as follows): . . . if you press the wrong key then a buzzer will sound. The time which it takes to complete this task is a good indication of intelligence; you should be able to solve it in two minutes."

At the conclusion of the experimental session, subjects of all groups were asked the following questions:—How did you go about the task of solving the problem? Did you change your method of solving the problem? Did the problem itself seem to change in any way? Did you suspect that the problem could not be solved at all? What do you consider the purpose of this experiment to be? What am I trying to find out?

III

Results

The effects of frustration on the learning time for the soluble problem.

Two subjects from the group undergoing 50 trials frustration without incentive, and three from the group undergoing 100 trials with incentive, withdrew from the experiment without having learnt the problem.

Positive skewness is evident in the distributions of the six groups, and in order to test hypotheses 1 and 2 an analysis of variance was carried out on the logarithmic transformation of the scores. This transformation was advisable since the data were of the same kind as might be obtained for a "survival" distribution and it might, therefore, be assumed that they were quasi-logarithmic in form. The results of this analysis of variance are given in Table II. Student's t tests for small samples were calculated for the differences between individual groups.

The comparisons of groups with 5-trials frustration and groups with 50- and 100-trials frustration shows an increase in the learning times for the soluble problem significant at 0·0005 level of confidence. The differences between groups with 50 trials and 100 trials are not significant and in one case are in the direction opposite to that put forward in hypothesis 1. Thus it was necessary to reject hypothesis 1 in its present form, there being no direct relation between the increase in learning time and the duration of the frustrating period.

Differences between the mean learning times of groups working under the conditions of "no incentive" and "incentive" all occurred in the direction predicted in

hypothesis 2, and the overall F ratio (Table II) shows that this effect is significant at the 0·05 level of confidence. Tests between the groups taken in pairs, however, are non-significant. It may be concluded, therefore, that the results give some support to hypothesis 2.

TABLE II

VARIANCE TABLE

Source	*Deviance*	*d.f.*	*Variance estimate*
Incentive	0·35838	1	0·35838
Frustration	5·79584	2	2·89792
Interaction	0·06125	2	0·03062
Individual differences	9·72901	138	0·07050
Totals	15·94447	144	

Interaction $F = 0·0434$			n.s.
Incentive $F = 5·084$	$n_1 = 1$	$n_2 = 138$	$P < 0·05$
Frustration $F = 41·105$	$n_1 = 2$	$n_2 = 138$	$P < 0·001$

The effects of frustration on the degree of stereotypy of behaviour.

While an increase in the learning time for the soluble problem may be thought of as an indirect measure of stereotypy and rigidity, a more direct method of measurement is possible. Responses of six trials or more to the same key were taken as evidence of stereotypy; in practice, such stereotypes ranged from 6 to 75 responses in duration. While this form of response was sometimes made in an attempt to solve the problem, an examination of the results and in particular of the verbal reports of subjects shows that in the majority of cases the subjects did not limit their responses in this way for rational motives and, indeed, were often unaware that they had been responding on only one key for any length of time.

Two comparisons were made between experimental groups, one between the mean duration of the stereotyped responses for each group, and the other between the proportion of stereotypes to the total number of responses made by members of each group. The second measure of stereotypy has the advantage of taking all the responses, stereotyped and non-stereotyped, into consideration, while the first only selects the stereotyped responses for comparison.

The F ratios for the effect of duration of frustration (Table III) on stereotypy are highly significant; and on the basis of these results individual t tests for the differences between groups were carried out.

TABLE III

F RATIOS FOR STEREOTYPY

A. Mean duration of stereotypes:			
Frustration $F = 21·853$	$n_1 = 2$	$n_2 = 138$	$P < 0·001$
Incentive $F = 0·643$			n.s.
B. Mean proportion of stereotyped responses to total number of responses.			
Frustration $F = 19·248$	$n_1 = 2$	$n_2 = 138$	$P < 0·001$
Incentive $F = 1·857$			n.s.

As with the learning times discussed above the differences were in all cases significant ($P < 0{\cdot}025$) between groups with 5 trials and groups with 50 or 100 trials frustration. Differences between groups with 50 and 100 trials frustration were, however, in both cases in the direction opposite to that put forward in the hypothesis. While stereotypy increases significantly with an increase in the duration of the period of frustration from 5 to 50 trials, a further increase to 100 trials does not lead to greater stereotypy, and the results suggest that under conditions of prolonged frustration the degree of stereotypy actually decreases. Hypothesis 3 cannot, therefore, be substantiated in its present form, there being no direct relation between the duration of the frustrating period and the degree of stereotypy of behaviour.

The differences between the incentive and non-incentive groups again occur in the direction predicted, but none of these differences is significant and no conclusive support can be given to hypothesis 4.

The relation between increased learning time for the soluble task and increased stereotypy of behaviour.

The increase in learning time and in the degree of stereotypy of behaviour under frustrating conditions were both used as measures of the disorganization of the subject's behaviour. In order to determine the agreement between these criteria, subjects were ranked according to their "scores" on both criteria and a correlation between these rankings was calculated using Spearman's "rho" statistic. Since "rho" in this case, is a measure of agreement between two rankings for which there is no known objective order the form "rho b" given in Kendall's (1948) treatise was used. A rank correlation value of $+0{\cdot}34$ was obtained between the variables of learning time and the proportion of stereotyped responses to other responses, a critical ratio of 4·0 affirming a high level of significance.

Vacillation of response.

The degree of vacillation of response was obtained through the records of the pressure made by the subjects on the response keys. The records of those subjects showing a high degree of continuous vacillation (approximately 41 per cent. of the total number of subjects) tend to reveal three phases:

(*a*) an initial steady period with little sign of vacillation;
(*b*) an intermediate period with fairly marked vacillation; and,
(*c*) a concluding period in which the amount of vacillation falls to zero.

There is, then, evidence of increasing vacillation during the frustrating situation, with a tendency for this vacillation to decrease as the subject gains mastery of the problem. It was also noticed that during periods in which stereotypes were formed, the records exhibited less vacillation, this effect demonstrating a possible "adjustive" value of the stereotype in reduction of tension as was suggested by Maier (1949) in his studies of stereotyped behaviour in the rat.

Analysis of the Verbal Reports.

The verbal reports of the subjects were found to be particularly revealing as to the existence of covert anxiety and tension; feelings of uneasiness and aggression are especially evident among those subjects undergoing the longer periods of frustration of 50 and 100 trials.

From the questions put to the subjects it would seem that the insoluble nature of the task was well disguised. None of the subjects thought the problem to be insoluble, though a few stated that at some time or other during the experimental

session they had thought that it might be too difficult for them to solve. Twenty-eight of the subjects suspected that the task had changed at some time during the experimental session, and a number thought that the change might have been made because of their inability to solve a difficult problem. In general, the method chosen to solve the problem was for the subject to test out a series of hypotheses as to what the actual pattern might be, abandoning each hypothesis as it proved false. The difficulties experienced by the subjects in the solution of the problem appear to have been, first that of finding a starting point to the pattern, and, second of deciding just how long the patterned sequence might be.

Direct questions were put to some of the subjects in an attempt to find out the reasons why they had assumed a "stereotyped" response. Replies, on the whole, were vague and unsatisfactory. Many of the subjects had no idea that they had been pressing down the same key for any length of time. Others stated that the time limit of five seconds for each response was too short for them to stop and think and this had probably led to their pressing down one key continuously. Subjects in Marquart's (1948) experiment who formed stereotypes reported that they felt compelled to continue using the same basis of choice. None of the subjects in the present experiment gave any direct hint of such compulsion, but the fact that the speed at which responses were demanded was found to be excessive by some subjects and that this was put forward as a reason for forming a stereotype may give some indirect evidence for the compulsiveness of the response. What is certain, is that for the majority of subjects the stereotype did not represent a rational attempt to solve the problem; usually the subject was unaware that he had shown any rigidity of behaviour, and when confronted with evidence of a stereotype was unable to give any explanation as to why it had occurred.

IV

Conclusions and Discussion

The results of the experiment show that an immediate disorganization of behaviour occurs as a result of frustration, this disorganization being shown by increase in the learning time of a subsequent problem and by an increase of stereotypy of behaviour both during and after a period of frustration by an insoluble problem. Under conditions of prolonged frustration, however, a point is reached after which further frustration does not lead to a further increase in the time taken to learn a subsequent problem or in the degree of stereotypy.

Differences between the groups working under conditions of "incentive" and "no incentive" in all cases show that stereotypy of behaviour is increased by the conditions of "incentive," but in two analyses out of three these differences were not statistically significant.

The assumption of the experimental method that the increase in learning time and in stereotypy were both measures of the disorganization of the behaviour of subjects was tested by means of Spearman's "rho" statistic. The value of $+ 0{\cdot}34$ was obtained when this correlation was carried out between the variables of learning time and the proportion of stereotyped responses to other responses. It is possible that this coefficient might have been considerably higher had more of the subjects shown evidence of stereotypy.

Finally, there was a definite trend towards increased vacillation over responses during that period of the task when frustration might be considered to have been at its peak, this vacillation being shown up by fluctuations in the pressure of the subject's hand on the response keys. It was also observed that vacillation did not

occur to the same degree during periods when the subject had formed a stereotype to one key, a fact which suggests that the stereotype represented some form of adjustment by the subject to the frustrating situation.

Hypotheses 1 and 3 that the learning time for the soluble task and the degree of stereotypy would increase in direct relation to the duration of frustration, were formulated largely on the basis of such experiments as those of Patrick (1934) on the disorganization of behaviour caused by an "emotional" response, and Maier (1949) and Marquart (1948) on frustration, all of whom found that behaviour under stress became increasingly more disorganized and stereotyped. The results of the present research, however, show that the relation between the duration of the period of frustration and the degree of disorganization of behaviour is not direct, and that under conditions of prolonged frustration the subject is likely to make some form of adaptation to the frustrating situation. From the study of the rat under conditions of frustration Knöpfelmacher (1953) has obtained similar results. Knöpfelmacher found that rats under conditions of high punishment during the frustrating situation showed a tendency to relinquish a stereotyped mode of behaviour more quickly, when a soluble problem was introduced, than those under low punishment. In his research the punishment consisted in preventing the animals from escaping from a water discrimination unit for periods of 8 seconds, for the low punishment groups, and 80 seconds for the high punishment groups.

Mackworth (1950) has also reported a recovery period in the efficiency of behaviour after a long period of stress. Mackworth found that subjects working under stress conditions produced by extreme temperatures or anoxia showed a gradual decline in the efficiency of their performance at certain visual and auditory tests, but after a period of between 20 and 30 minutes a recovery period was evident during which efficiency of work increased.

Maier (1949) claims that as a result of his experimental work on rats, there is good reason to suppose that "motivated" behaviour differs in *kind* from "frustration-instigated" behaviour. Working on human beings in a frustrating situation, Marquart (1948) states that the distribution of results from her experiments show definite signs of a bimodality which would be required to support Maier's hypothesis. In the present investigation no evidence to support such a distinction could be found. An examination of the distribution curves of the "scores" for learning time and degree of stereotypy showed no signs of bimodality and there were no other qualitative distinctions observable between the behaviour of the subjects under conditions of increasing frustration. Instead of considering behavioural deviations to be different in kind from "normal" motivated behaviour, it is suggested that behavioural deviations, like other reactions of the body to stress, undergo three stages analagous to those of the general adaptation syndrome put forward by Selye (1950) in the study of psychosomatic conditions.

(1) An alarm reaction, during which behavioural deviations increase progressively;
(2) A stage of resistance, in which there is some adjustment to the frustrating situation;
(3) A stage of exhaustion, in which the subject shows extreme fatigue or aggression, or withdraws from the problem.

When Cannon (1932) formulated his "emergency reaction" he showed the utility of a physiological mechanism for the explanation of behaviour under conditions of immediate stress. Selye's general adaptation syndrome is an extension of Cannon's emergency reaction to situations of prolonged stress and may well prove to be the best theoretical basis for behavioural studies of stress, frustration and fatigue.

This paper is an abridged version of a thesis—"The Effects of Frustration on Learning in Human Subjects"—submitted and approved for the Degree of Doctor of Philosophy in the Faculty of Science, University of London, 1952. The author would like to express his appreciation and gratitude to Professor R. W. Russell who supervised the research.

References

CANNON, W. B. (1932). *The Wisdom of the Body.* New York.

JONES, L. C. T. (1952). The Effects of Frustration on Learning in Human Subjects. Ph.D. Thesis, University of London Library.

KENDALL, M. G. (1948). *Rank Correlation Methods.* London.

KNÖPFELMACHER, F. (1953). The rigidity of position stereotypes in the white rat under varying conditions of reward and punishment. Ph.D. Thesis, University of London Library.

MACKWORTH, N. H. (1950). *Researches on the Measurement of Human Performance.* London: H.M. Stationery Office.

MAIER, N. R. F. (1949). *Frustration—the study of behaviour without a goal.* London.

MARQUART, D. I. (1948). The pattern of punishment and its relation to abnormal fixation in adult human subjects. *J. gen. Psychol.*, **39**, 107–144.

PATRICK, J. K. (1934). Studies in rational behaviour in emotional excitement: II, The effect of emotional excitement on rational behaviour of human subjects. *J. comp. Psychol.*, **18**, 153–195.

SELYE, H. (1950). Adaptive reactions to stress. In: *Life Stress and Bodily Disease*, ed. Wolff, H. G., Proceedings of the Assocn. for research in Nervous and Mental Diseases, **29**.

UNDERWOOD, B. J. (1949). *Experimental Psychology.* New York.

CAN A TRAINED SUBJECT JUDGE HIS AUDITORY SENSITIVITY?

BY

MICHAEL WERTHEIMER

From the Psychological Laboratory, Wesleyan University, Middletown, Conn.

Auditory thresholds were measured daily on three subjects over an extended period of time, and the subjects postdicted whether the level and variability of the day's measured thresholds were higher or lower than on the preceding day, without their having knowledge of their performance. It was found that their judgments agreed with the obtained psycho-physical data better than would have been expected on the basis of chance alone. These results raise the question of the possibility that a subject's understanding of, or hypothesis concerning, the experimental outcome may influence his behaviour.

While a negative answer to the question which is the title of this note might lead an investigator in the field of sensation to breathe more easily, an answer in the affirmative would raise quite serious problems. In an attempt to answer the question experimentally, absolute auditory thresholds for clicks were measured monaurally on three subjects daily at the same time of day for 46 consecutive days. As soon as a subject's threshold had been measured, but before he saw what his day's performance had been, he tried to estimate his performance relative to his performance on the previous day. These "postdictions" were made both with respect to the value of the absolute threshold, i.e., whether it was higher or lower than on the previous day, and with respect to the variability of response, i.e., whether the standard deviation of the six samples taken daily was larger or smaller than on the previous day.

These experiments were performed in the spring of 1950 at the Harvard Psycho-Acoustic Laboratory; a detailed account of the procedure has been given elsewhere (Wertheimer, 1953). All three subjects had had extensive prior experience as subjects in auditory psycho-physical experiments.

TABLE I

Relation Between True and Estimated Levels of the Absolute Auditory Threshold for Clicks

True	*Judged*	*Subject D*	*Subject N*	*Subject M*	*Total*
H	H	12	8	16	36
H	L	4	10	5	19
L	H	4	1	8	13
L	L	13	13	12	38
Total		33	32	41	106

Tables 1 and 2 show the results of the analysis of the post-diction data. Entered in the tables are the number of times that each subject gave a judgment of "higher" or "lower" when the actual measurements indicated that the threshold or variability were higher or lower. For Table I, if the threshold was higher than on the previous day, and the subject postdicted "higher," HH is entered; if the threshold was higher, and the subject postdicted that it was lower, HL is entered; LH and LL have comparable meanings. Table II is constructed in the same way, except that "higher" and

"lower" refer to the standard deviation of the responses. On occasion the subject said, "The same as yesterday," or "I don't know;" therefore the vertical total of the table columns for each subject is less than the 45 possible.

TABLE II

Relation Between True and Estimated Levels of the Standard Deviation of Threshold Measurements

True	*Judged*	*Subject D*	*Subject N*	*Subject M*	*Total*
H	H	5	12	6	23
H	L	2	2	6	10
L	H	4	6	4	14
L	L	3	5	13	21
Total		14	25	29	68

Table I shows that 74 times out of 106, the subjects' estimates of the level of the threshold were correct, while they were incorrect only 32 times. The estimates of variability in Table II similarly were correct 44 times out of the total of 68 postdictions, incorrect only 24 times. In the case of both tables, the subjects' combined estimates of their performance differed from chance at a significance beyond the 1 per cent. level, by testing the difference from chance against the standard error of the proportion.

The subjects felt, introspectively, that they based their postdictions on many different kinds of cues, such as how loud the clicks seemed to be at and above threshold, what the perceptual quality of the clicks was, how often click quality changed during a measurement, how attentive the subject felt, and what the fluctuations in attention seemed to be.

Whatever the bases for these estimates, the fact that a subject has some ability to tell how he is doing in such an experiment raises some interesting questions, including the possibility of conscious or unconscious intent if he has some knowledge or hypothesis about the design or outcome of the experiment. It is, for example, not inconceivable that a subject's responses would be different if he expects the quantal hypotheses to hold, from what they would be if he believes the phi-gamma curve to be a more adequate description of the threshold psychometric function. In studies of cyclical changes in threshold, a subject may unwittingly be producing some regular cycles in his response sequence if he believes that cycles exist, while such cycles might not occur if the subject is convinced that threshold variations are random.

It would seem, then, that the use of a subject who is sophisticated about sensory research and theory would be inadvisable. Yet most threshold research requires a subject who is highly trained in the discriminations required for precise determinations. It is, unfortunately, rare to find a subject experienced in one without his being experienced in the other. Although it may be possible to use naive but well-trained discriminators as subjects in an occasional research design, in the overwhelming majority of contemporary studies, the investigator using trained subjects in psychophysical studies would do well to keep in mind the possibility that the subject's understanding of the experiment may not be unrelated to the results produced.

Reference

Wertheimer, M. (1953). An investigation of the "randomness" of threshold measurements. *J. exp. Psychol.*, **45**, 294–303.

DISCRIMINATION OF SIZE BY SIGHT AND TOUCH

BY

R. P. KELVIN

From the Psychological Laboratory, University College, London.

The experiment was designed to discover whether accuracy of discrimination differs when the discrimination is made in only one sense from when it is made across two senses. The evidence produced by the experiment does not warrant one to conclude that such a difference exists. The absence of such differences has been considered in relation to more general perceptual problems.

Introduction

This paper reports the results of an investigation into the accuracy with which human beings can make discriminations in size between two objects of which one is seen and the other is held in the hand. There are at least two good reasons for considering discriminations which involve more than one sense. In the first place almost any information about the sensory capacity of normal people may turn out to be a useful guide in the clinical assessment of injury and impairment. The second ground for interest is more general. There are some physical characteristics of objects which may be apprehended by more than one sense. For instance the size of an article may be perceived both by sight and by touch. Any theory about the perceptual organization and processes of either of these two senses by itself will be incomplete unless it also indicates how they can collaborate. Conversely, some helpful knowledge may be gained about each sense by studying instances of such collaboration.

The word "sight" rather than the more technical term "vision" has been used throughout this paper to emphasize that the main purpose of this study was the accuracy of discrimination in the everyday activity of looking at (divers) objects. Similarly "touch" has been used to refer to the everyday activity of perceiving things by touching them, feeling them and holding them in the hand. In this way "touch" includes both tactile and kinaesthetic processes. These definitions of "sight" and "touch" distinguish perhaps rather sharply between the performance of the subject and the physiological processes which subserve that performance. The distinction is deliberate, because there is no intention to investigate here the central or peripheral mechanisms which mediate perception with either or both senses.

Very little work has been done on discrimination across two senses. What has been done has concerned itself primarily with the P.S.E., that is, with whether things perceived by touch appear larger or smaller than things seen (Jastrow, 1889). Such work is not really relevant to the present inquiry which concerns *accuracy* of discrimination. It seems, therefore, most useful to state at once the problems of this experiment.

Aims and Basic Assumptions

The general aim of this paper was to discover whether the degree of accuracy of size discrimination across the two senses of sight and touch differed from the degree of accuracy achieved by each sense when it acted on its own.

Accuracy was held to be most usefully assessed on the basis of the Standard Deviation of two categories of judgment obtained under a Constant Method procedure.

The overall problem had to be solved by means of a number of subsidiary and more specific investigations:

(*a*) Accuracy of discrimination had to be determined for sight and touch, each acting by itself.

(*b*) Since the receptors (hand or eye) are available on both sides of the body, and since they can function either singly or in conjunction, the possible effects of laterality had to be kept in mind.

(*c*) In cross-modal discriminations, the results might be disproportionately affected either by the sense which perceives the standard or the sense which perceives the comparisons. Therefore both senses must act in both capacities.

In order to get some coherent pattern out of these second-order investigations the following assumptions were made:

(*a*) Discrimination in human beings requires the efficient working of both peripheral receptors and central nervous mechanisms.

(*b*) Each system of receptors is connected with at least one cortical centre. Cortical centres are connected with one another.

(*c*) *Accuracy* of discrimination decreases with every increase in the number of neural centres which may be assumed to subserve the discriminatory processes.

It is important to realize that these basic assumptions simply provide the framework within which the particular hypotheses of the experiment can be formulated. They are deliberately naive physiologically. They have been stated like this primarily to avoid begging questions concerning lateral or sensory dominance. Neither one side of the body, nor one sense is favoured by these assumptions. If such dominance exists, the fate of the hypotheses to which these assumptions give rise would display it. For these assumptions lead to the prediction that there will be three levels of accuracy arranged in a hierarchical order. Position in the hierarchy is determined solely by the relative simplicity or complexity of a given combination of receptors.

(1) The best performance would be predicted for the condition in which the same receptor (e.g. the right hand) perceived both the standard and the comparisons.

(2) Next in order would be the condition in which two stages were involved in the perception; i.e. where the standard and the comparisons were presented on the same side of the body, but to different senses, or where the standard and the comparisons were presented to the same sense but on opposite sides of the body.

(3) The condition for which least accuracy is predicted would be that in which the standard and the comparisons are presented to different senses and to opposite sides of the body (e.g. to the left hand and the right eye).

Two perceptual conditions are not covered by these three levels. It is not possible to predict how accuracy of discriminations involving both sides of the body simultaneously will compare with the degree of accuracy achieved by one side acting alone. However, it is possible to predict that in such bilateral discriminations, cross-modal discrimination will be cruder than discrimination within the same sense.

These three predicted levels give rise to two classes of hypotheses.

Class I: These hypotheses stipulate that there will be differences between different levels of receptor arrangements. They predict that the mean S.D. for the first level will be smaller than that for the second, and that the mean S.D. for the third level will be larger than that for the first two. Also within this class comes one prediction of *no* difference between contra-lateral comparisons in the same sense and ipselateral comparisons using both senses (these are the two types of condition that separately make up the predicted second level of accuracy).

This is the class of hypotheses which is of primary psychological interest. The results of the tests of these hypotheses can act as a basis for comparing cross-modal discrimination with the discriminatory capacity of one sense.

Class II: The hypotheses in this class are more statistically necessary than psychologically significant. They are all Null Hypotheses stipulating that there is *no* difference between the mean S.D. for any of the types of conditions which make up a given level of complexity.

In order of statistical procedure Class II hypotheses must be tested before those of Class I, since the latter can only be tested after it has been shown that each level of complexity is itself homogeneous.

It is not necessary to go into the details of the constituent hypotheses of each class. There were seven hypotheses of Class I and six of Class II. It was decided to reject a Null Hypothesis if F exceeded P = 0·05. It was also decided to accept as significantly different the difference of means where t exceed P = 0·05.

Design

(Throughout the remainder of this paper, the sense perceiving the standard will be written first, that perceiving the comparisons will be placed second.)

1. There are four possible combinations of receptors in terms of modalities:

(*a*) Touch—Touch (*b*) Touch—Sight
(*c*) Sight—Touch (*d*) Sight—Sight.

2. There are five possible combinations of receptors in terms of the side of the body on which those receptors may be located:

(*a*) Left—Left (*b*) Left—Right
(*c*) Right—Right (*d*) Right—Left
(*e*) Both—Both.

Each combination of modalities may be investigated with every combination of laterality. There are, therefore, $4 \times 5 = 20$ conditions to be investigated for variations in degree of accuracy.

These investigations were carried out with five subjects (2 males, 3 females) being used for each condition (cell). Each subject was tested for the condition of one particular cell only, and only one test was made with each subject. On the assumption that the subjects formed a random sample, the design permitted a straightforward application of the analysis of variance, with a minimum number of the complications that tend to confuse this procedure in psychological investigations, especially where repeated tests are made on the same subject.

Apparatus

The Stimuli:

The stimuli were a series of discs, made from aluminium 3·5 mm. thick. Their diameters ranged from 57·5 mm. to 62·5 mm., in step intervals of 0·5 mm. The size of the standard disc was 60 mm., and there were two discs of that diameter.

The weight of each disc was made to be equal to the weight of the smallest disc. This was done by scooping out a thin layer of metal from one surface of each of the larger discs. The amount of metal scooped out of any one disc was equivalent to the amount of metal by which that disc was larger than the smallest disc.

A thin strip of steel was fitted on the scooped out surface of each disc.

The Setting:

(*a*) For presentation to sight: The discs were placed on a wooden board which had been covered with green baize. This board was placed before the subject in the frontal-parallel plane, about 1 metre from the subject's eyes.

The board measured 46·5 × 31 cm. A magnet was fitted on the back of the board, 10·5 cm. from the top, and 15·5 cm. from the side of the board. Every disc could be held on on to the board by placing it so that the steel strip faced the magnet which was covered by a thin layer of baize only. Two such boards were made to enable easy presentation of the standard and the comparison stimuli in some of the visual experiments.

(*b*) For presentation to touch: The subject had to put his hand through a screen, 58·5 cm. × 33 cm. with an opening of 45 cm. × 11 cm. at its base. This opening was sufficiently wide to allow ample movement of the hands and lower arms. The height of this screen was low enough to enable the subject to have a clear view of the discs on the boards whenever this was necessary. In order to keep the general visual field as uniform as possible, this screen was also covered with green baize.

On the back of this screen there were fixtures for a light and a thin metal rod.

The light was used in the visual tasks. It illuminated the discs on the boards without casting any noticeable shadows.

The rod was fitted so that it projected above the middle of the screen. When the rod was in position it acted as the axis for a swivelling occluder. This occluder consisted of two boards, 30 × 23 cm. fixed at 90° to each other on a thin wooden block. They were covered with green baize. A hole was drilled into the block. The occluder could thus be made to fit over and pivot on the rod projecting above the screen.

By using the occluder it was possible to prevent the subject from seeing either the right or the left, or even both boards, e.g.:

(1) Suppose that the subject has to make comparisons by using his left eye only: he would be seated in such a way that the occluder obscured all views of the boards from his right eye.

(2) Suppose now that the subject had to make comparisons between his right and his left eye: he would first be shown the standard on his left, the occluder shutting off the right eye; the occluder would next be moved 45° and thus obscure all view of the boards for both eyes; it would then be swung completely over and so obscure the view for the left eye while allowing the right eye to see the disc.

By using this device, it was possible to reduce to a minimum the fatigue which comes either from the wearing of eyeshields or from the repeated opening and closing of one or the other of the two eyes. It also reduced to a minimum the optical distortions due to changes in light adaptation (Horowitz, 1949).

Procedure

Presentation of the stimuli:

The stimuli were presented in accordance with a Constant Method Procedure. The standard stimulus had a diameter of 60 mm., and there were four larger and four smaller discs for comparison. These comparison discs measured 57·5 mm., 58·5 mm., 59 mm., 59·5 mm., 60·5 mm., 61 mm., 61·5 mm., 62·5 mm. in diameter; there was also a ninth disc of 60 mm. to match the standard. Of the possible stimulus series the discs with diameters 58 mm. and 62 mm. were omitted. In most cases during pilot studies during which these discs were included, hardly any errors were made with them. Cutting these discs out considerably relieved the tedium of the experiment for the subjects. The majority of the subjects of the pilot investigation became resentful and bored by the continued re-presentation of stimuli which for the most part were obviously ("pointless") different. But as a check the penultimate rather than the extreme members of the series, 57·5 mm. and 62·5 mm., were cut out, to allow linear interpolation between the third and fifth value at each end of the series if necessary.

Each of the comparison stimuli—including the disc which matched the standard for size—was presented six times for comparison with the standard. Each of the nine comparison stimuli was presented once before any one of them was presented a second time, and so on for the second and third presentations, etc.

The order of presentation was random for the first three columns; the second three columns were mirror images of the first.

The 60 mm. disc (i.e. the standard size) was presented alternately to the receptor for the standard and to the comparing receptor.

The stimulus for the standard receptor was always presented first, for approximately 3 seconds. There was then a pause of about 6 seconds. After that the subject was given the stimulus for the comparing receptor for a period of about 3 seconds, and asked to judge this second stimulus either larger or smaller than the first. The judgments equal or doubtful were not allowed. After each judgment there was a pause of approximately 6 seconds during which the experimenter made notes of the subject's answer and selected the next two comparisons.

These times are approximations. The experimenter had a stopwatch, and in the large majority of cases the timing was adhered to very well. Occasionally, however,

TABLE I

STANDARD DEVIATION; 20 CONDITIONS, 5 SUBJECTS PER CELL

Standard Comparison	*Left Left*	*Left Right*	*Right Right*	*Right Left*	*Bilateral*
Touch Touch	1·22	1·37	0·69	1·46	1·28
	1·14	1·13	1·07	1·37	0·77
	0·98	1·18	1·23	1·45	1·84
	1·49	0·79	0·48	1·12	1·11
	1·27	1·21	1·08	1·44	1·23
Mean	1·22	1·136	0·91	1·304	1·246
Touch Sight	1·38	1·21	0·86	1·23	1·18
	1·25	1·75	1·62	1·14	0·86
	1·46	1·61	1·53	1·68	1·04
	1·56	1·79	2·10	1·48	0·90
	1·14	1·02	1·09	0·76	0·88
Mean	1·358	1·476	1·44	1·258	0·972
Sight Touch	1·34	1·44	1·14	0·85	1·31
	0·87	1·14	1·25	0·58	1·01
	1·00	1·25	1·46	1·43	0·94
	1·17	1·31	1·23	1·29	1·52
	0·85	1·17	1·07	1·43	1·11
Mean	1·126	1·262	1·23	1·116	1·178
Sight Sight	0·99	1·65	1·38	1·40	0·79
	1·02	1·35	1·55	1·27	1·19
	0·29	1·34	0·83	1·69	1·06
	0·35	1·40	1·11	1·02	1·07
	0·89	0·90	1·21	0·92	0·87
Mean	0·708	1·328	1·216	1·26	0·996

there was a slip: the stimulus disc might drop from the display-boards, the subject sometimes gave his answer quickly, at other times he would lag a little; these, and similar happenings make it impossible to vouchsafe exact timing. In the main, nevertheless, the time was kept to much more closely than the experimenter had expected. Each session with any one subject lasted about 20 to 25 minutes.

Subjects:

Forty men and 60 women acted as subjects in this experiment. The subjects were volunteers and were told before volunteering the general aims of the experiment. They were none of them acquainted with the aims, methods and complications of experimental psychology. Their performance was thus not biased by preconceived notions about what happens in psycho-physical experiments. The men were mainly office staff at every level of competence, though there were some research scientists among them as well. The women were predominantly clerical workers and secretaries, but there were also some technicians and research workers among them. The subjects were taken singly, and the instructions varied with the twenty conditions under any one of which a given subject might be tested. However, all instructions stressed that the subject should attend as carefully as possible to his first impression of the stimulus; he was warned that trying to "work out" and thinking hard about the size of a disc would only make it more difficult for him to come to a decision; he was assured that he would do best by attending to first impressions.

The majority of subjects were right-handed. Those who reported that they were left handed were used for bilateral tasks only.

Results

Table I shows the S.D. for each subject, and the mean S.D. for each cell. The first two scores in each cell are the scores for the men in that cell, the remaining three values are the S.D.s of the women. Each value was computed by the Spearman summation method (*cf.* Woodworth, 1950, p. 403).

TABLE II

	Sum of Squares	*d.f.*	*Mean Square*	*V.R.1.*	*V.R.2*
Sides	6051·9	4	1512·975	1·909	1·979
Senses	5207·8	3	1735·9	2·191	2·271
Sex	0·167	1	0·167	0·211	0·218
Sides × Sex	2169·683	4	524·421	0·662	0·686
Senses × Sex	5954·7	3	1984·9	2·505	2·597
Sides × Senses	20436·1	12	1703·008	2·149*	2·23*
Sides × Senses × Sex ..	7497·32	12	624·78	0·789	0·764
Error	47535·33	60	792·25		
Total	94853	99			

Where the value in this table or subsequently is marked thus *, P = <0·05 and where ** P = <0·01.

Table II shows the three-way analysis of variance for this data. Since the triple interaction was not significant, its sum of squares was combined with that for the error term, and the resulting new mean square was used to obtain V.R.2 with 72 degrees of freedom. As the three-way analysis shows significant interaction between sides and senses, it was reasonable to proceed to test the various individual hypotheses to see whether there was a coherent pattern of differences of accuracy. It is only necessary to summarize the outcome of these tests.

All the hypotheses of Class II (i.e. all those stipulating *no* differences between the means of certain groupings of cells) were acceptable apart from one. The unacceptable hypothesis predicted that there would be no difference between any of the perceptual conditions in which the standard and the comparisons were presented to the same receptor (hand or eye). In other words the Null Hypothesis underlying that condition for which greatest accuracy had been predicted, had to be rejected. Consequently those hypotheses of Class I which were designed to test the relative degrees of accuracy of other perceptual conditions with respect to this highest level, could not be tested. However, when the remaining hypotheses of Class I were tested, none showed the predicted differences—or indeed any significant differences. The only hypothesis of this class which was acceptable was also the only one in this class which predicted that there would be no difference between the means of the two types of condition which make up the second level of accuracy (on the basis of the underlying assumptions of the experiment).

As a result of this failure to find any of the predicted differences between the various groups of cells tested, the cells were regrouped and tested for effects of lateral or sensory dominance.

Lateral Dominance: All the hypotheses were null hypotheses. All were acceptable, with one exception. The hypothesis stipulating that there would be no difference within the conditions in which the standard was presented to the left receptor, and where the comparisons were made in the same sense, had to be rejected. This meant that no test could be made between this group of cells, and the corresponding groups on the right side of the body. It was however plain that laterality did not affect crossmodal performance.

Sensory Dominance: The aim here was to test whether cross-modal performance was cruder than ipsemodal performance (i.e. performance in which the standard and the comparisons were perceived by the same modality). Each of the four conditions in terms of modalities was taken as a whole and tested for homogeneity. All but that in which both the standard and the comparisons were presented to sight, showed homogeneity. The sight-sight group had an $F = 4{\cdot}02$ for 4 and 20 d.f., which is significant beyond $P = 0{\cdot}01$. It was therefore not reasonable to proceed to test the difference between the ipsemodal and the crossmodal tasks. As a matter of interest the difference between the means of the two cross-modal tasks was tested. $t = 1{\cdot}409$ (d.f. $= 48$) which is not significant.

Both the Null Hypotheses based on the original assumptions, and the *post-hoc* hypotheses concerning lateral and sensory dominance thus fail to show systematic differences in the degree of accuracy achieved under the various conditions of the experiment. At the same time some hypotheses could not be tested because one of the groups between which differences were to be tested showed itself to be internally significantly inconsistent. On examining the raw scores, these internally inconsistent groups always contained the bottom left-hand cell of Table I: the condition in which the left eye perceived both the standard and the comparisons. Inspection of the values of this cell will show that it contains particularly low values for the standard deviation. This conglomeration of small S.D.s in this one cell is almost certainly responsible for the heterogeneity found within the groupings of cells in which this cell participated.

In the three-way analysis of variance, however, there is a noticeably small sum of squares for sex. So much is this the case that it was thought justifiable to examine the results of the men and the women separately in order to discover whether the results for the whole data have been obscured by a counterbalancing of the results of men and women respectively. Therefore all the hypotheses which had been formulated for the combined data were now tested separately for men and women. Very little more positive information was gained.

Men: The two-way analysis of variance showed no significant effects. The highest V.R. was for Sides (2·067, d.f. 4 and 20) and even that was not significant.

All the hypotheses of Class II were acceptable, and therefore all the hypotheses of Class I could be tested. Two of these were acceptable. One was the only null hypothesis in this class. The other showed that performance in the same sense but across opposite sides of the body was less accurate than performance in which the same receptor perceived both standard and comparisons.

There were no significant differences due either to lateral or sensory dominance. Again all hypotheses could be, and were, tested.

Women: The two-way analysis of variance showed a significant effect with respect to the Senses. (V.R. $= 3{\cdot}423$,* d.f. $= 3$ and 40.)

Again all hypotheses of Class II were acceptable and all the hypotheses of Class I

could therefore be tested. Three of these were acceptable. One was the null hypothesis of this class. Of the other two

(*a*) Cross-modal accuracy was found to be significantly cruder than ipsemodal accuracy, when the comparisons were presented on the same side of the body.

(*b*) Cross-modal tasks involving opposite sides of the body were found to be significantly less accurate than ipsemodal task performed ipselaterally.

There was no indication of lateral dominance for women.

There was, however, a significant difference between all ipsemodal means and all cross-modal means.

$$\left.\begin{array}{rl}\text{Ipsemodal Mean} & : \ 1{\cdot}084 \\ \text{Cross-modal Mean} & : \ 1{\cdot}289\end{array}\right\} \ t = 2{\cdot}54^{**} \ \text{d.f.} = 58$$

It is, therefore, clear that a slightly more articulated picture emerges when the results for men and women are separated. There is, for instance some suggestion that the performance of men is more affected by the sides of the body involved in the discrimination, whereas the performance of women seems to be more affected by whether one or two senses have to be used. However, though the results for the separated data may seem slightly more positive, this does not mean that it is either justifiable, or even useful to treat the results of men and women separately. It should be pointed out that the age range of the men and the women differed significantly, but that there was no significant correlation between age and S.D.

Discussion

The first problem which arises is to decide whether it is more reasonable to interpret the results for men and women combined, or whether to follow up the results of the two sexes separately. The results for the combined and separated data are sufficiently divergent for it to be impossible to accept both. Thus some hypotheses which had to be rejected for the combined results were acceptable for the separated data. The writer has come to the conclusion that on the basis of the information available to him there is not sufficient cause to treat the results for men and women separately.

From a purely statistical standpoint there were originally thirteen hypotheses. Six of these postulated that there would be differences between the conditions tested under them. Seven of them (including one from Class I) predicted that there would be no differences among the conditions tested under them. For the combined results six out of the seven hypotheses predicting homogeneity could be accepted; only one had to be rejected. Three only of the hypotheses which stipulated differences could be tested; not one was acceptable. The combined results thus certainly show no strong tendency for accuracy of discrimination to vary systematically with the conditions under which the discrimination has to be made. This same conclusion follows from the tests of the *post-hoc* hypotheses concerning lateral and sensory dominance.

The position is similar for the separated data. Because the numbers per cell were reduced, the hypotheses covering bilateral discrimination were telescoped (i.e. all the four possible bilateral conditions were taken together and tested for homogeneity, and were found homogeneous). There were, therefore, only six hypotheses postulating homogeneity and only five predicting differences. Both for men and for women all the hypotheses about homogeneity were acceptable. For men only one, for women only two of the hypotheses concerning differences could be accepted; and it should be noted that all these hypotheses were testable. Moreover, of this group of hypotheses, those acceptable for men and those acceptable for

women were different. While at first glance this fact would support the policy of keeping the results for men and women apart, the writer believes that such a policy would be mistaken, or at least misleading.

An overall examination of the results, both combined and separated, emphasises the absence of differences rather than their systematic distribution. In the case of the separated data a total of twenty-two hypotheses was tested only three of which showed differences to exist. In view of the small number of cases per cell, and the strong indication of homogeneity over the whole table, it seems reasonable to suggest that these differences were only accidentally significant. Also, since the left eye failed to predominate in the separated data, its effect on the combined results was probably due to an accidental grouping of small S.D.s in one cell rather than to a systematic difference made more apparent by a larger number of cases.

In other words the writer feels that there is not sufficient evidence to suppose that the twenty conditions examined in this experiment show any coherent pattern of differences. If the various hypotheses had been acceptable, or if at least some of them had shown some ordered differences in accuracy between various conditions, then the small number of cases tested under each condition would have been compensated for by the consistency of the results. But one cannot make generalizations on the basis of five subjects. At the most one can state that in this experiment one or two conditions did not produce results similar to the rest. Such a statement would be an *ad hoc* comment on the experiment, with no consequences. It would in fact be tantamount to the admission that those differences might be due to sampling, and that would bring one straight back to the position the writer has chosen to adopt. However, the conditions for which differences have been found must be investigated further, so that if the differences are confirmed, they can be put forward as generalizations rather than as asides.

These results and conclusions are not really very surprising. Certainly tradition tends to favour the prejudices concerning handedness; that most persons are either right-handed or left-handed. Yet both observational and experimental investigations show that there is no simple form of "handedness." Whether in terms of the hand used in everyday activities, or whether in terms of special laboratory tasks, authors have been repeatedly forced to the conclusion that handedness must be defined in relation to a particular situation. (Cuff, 1931; Downey, 1933.)

The evidence for eye-preference is no stronger. Buxton and Crossland (1937) report that eyedness tends to vary with the test which is used to establish it. Moreover, since the eye transmits to both hemispheres it would in any case be unwise to generalize from eye-preference to the dominance of one hemisphere. There is also some evidence that the preferred eye is not necessarily the one with greatest acuity (Gahagan, 1933).

The absence of lateral dominance in the present experiment is therefore not surprising. This absence of evidence for lateral dominance only partly explains the absence of difference between ipselateral and contralateral comparisons. Though it may be permissible to infer that either side of the body is equally sensitive, one cannot therefore assume that there will be no difference between ipselateral and contralateral comparisons; it might, thus, be quite reasonable to assume that the two hemispheres and receptor mechanisms can act equally well, but that they have only rudimentary connections. However, anatomical connections are known to exist. Further, the recovery of functions after brain-lesions, both in man and animals, and the many forms of bilateral transfer, make it clear that in many discriminatory processes both hemispheres are at least partly involved; even where at first sight only one hemisphere would need to be active.

The issues raised by crossmodal discrimination are more complicated. The basic assumptions led to the prediction that crossmodal accuracy would be cruder than ipsemodal accuracy. The results for women showed that this prediction may yet be found tenable, although it had to be rejected for the combined data. But whether cross-modal discrimination turns out to be cruder than ipsemodal discrimination or not, it is quite obvious that human beings can estimate differences of size between a visually and a tactually presented object to a remarkable degree of accuracy. Speaking loosely, on the basis of the present experiment it seems reasonable to suggest that a difference of 1·5 mm. in diameter for objects varying around 60 mm. in diameter, would be detected quite easily. It has, however, almost been a maxim of experimental psychology that estimation of size is dependent on a host of cues which are ostensibly unrelated to the stimulus being judged. This belief is well illustrated by Graham (1950). Writing about experiments concerning visual perception of size, he remarks:

> "Size matches are dependent on the presence or absence of stimuli other than those which may be considered to be *the* discriminative stimuli."

These "other stimuli" are generally thought to be stimuli derived from the environment in which "the stimuli" are presented. But in view of the degree of accuracy of the cross-modal matches of the present experiment, Graham's conclusions become a little doubtful. For while it is, prima facie, reasonable to admit the influence of the background on *visually* perceived size, there is no such background to the *tactually* perceived stimulus. It is therefore difficult to explain how the combination of background and stimulus in the case of sight, produces a visual impression of size which can be matched so accurately with a tactual impression of size, which is not affected by such an extraneous factor in the perceptual situation as the concomitant background. Though in the present experiment the writer took every precaution not to introduce disturbing elements into his visual presentations, he finds it hard to believe that he eliminated all the effects that are usually attributed to the background. He finds it even harder to believe that by some fortunate accident he presented his visual stimuli against a background, and at a distance, which yielded a visual impression of size that almost precisely matched the tactual impression of size in which such a background was lacking. Instead he ventures to suggest that his results would not have been very different for a considerable range of distances and with a fair amount of variation in the area of the display-boards. This may have to be investigated experimentally, but the writer feels sure that over a middle range of distances, and in the absence of distortions deliberately introduced into the background, the results of cross-modal matches would not differ significantly from the ones which he obtained. There are two consequences to this conclusion.

In the first place, it may be worth considering cross-modal matches as a special case of constancy phenomena. For a match across two senses demands that the stimulus patterns of the two modalities have some common features; and these common features can only arise from the characteristics of the stimulus object. In view of the absence of possibly similar effects of the background on sight and touch, these common features can only be certain physical characteristics of the stimulus. The perceptual situation would therefore seem to constitute an instance of "phenomenal regression" to the "real" object.

The second consequence is concomitant with the first. It is not sufficient to maintain that "constancy" effects themselves are due to cues derived from the experimental setting. Though environmental factors may play an important part in some visual constancies, they cannot account for the form of constancy found in

cross-modal comparisons. Constancy within one modality is determined by the relative consistency of, e.g. judgment of size, for conditions which are considerably different in terms of the physical effects of the stimulus on the organism. This consistency is mostly attributed to cues from the environment. These are thought to supplement those of the primary stimulus in such a way that they compensate the changing effects of that primary stimulus due to changes in, e.g. distance. In terms of Gestalt Psychology one might express this by saying that the field, composed of the stimulus and its background, is structured so that changes in one of its aspects tend to be counteracted by simultaneous changes in the remainder of the field. But while this balancing act of the perceptual mechanism may account for the consistency of judgments in one modality, it is difficult to see how it can account for the agreement in judgment of size between sight and touch. There is no common physical field or background to the visual and tactual stimuli used in this experiment which is analogous to the field or the rôle of the background usually postulated for results showing constancy. The problem is therefore to explain how the organism achieves such a high degree of accuracy in comparing across two senses. For such matches must be considered to depend on the physical characteristics of the stimuli judged and not on the structure of the background against which they appear.

The fundamental point of interest is thus not so much how favourably ipsemodal accuracy compares with cross-modal accuracy, but that cross-modal judgments can be made at all and relatively so accurately. The writer does not begin to know what the answer is. Only certain considerations suggest themselves.

Performances involving two modalities are not an experimental artefact but an everyday occurrence. There is virtually no activity in which there is "feedback" where this feedback does not come by means of a sense other than that used in the performance. This is true of mirror-drawing, playing cricket and even talking. To ask a person to make a discrimination across two senses is not therefore asking anything superhuman. The neural mechanisms may scarcely be known, but it must be assumed that they are so constituted that the specific character of each sense does not dominate over whatever it may be that, neurologically, the senses have in common.

It is quite clear that certain characteristics of objects may be apprehended by more than one sense. This is particularly the case with sight and touch, where the same feature of the stimulus can be perceived by both senses. But each sense apprehends this feature in its own way. The writer wishes to suggest that an amended concept of the "Schema" as propounded by Head (1920) may be helpful in thinking about cross-modal phenomena. On the basis of his clinical observations Head found himself forced to conclude that there is no fixed image of the body which could act as a standard for the appreciation of postural change. Posture was continually changing in some way or other. Normal human beings nevertheless had the capacity to indicate, for example, where a certain movement had deposited their arm with respect to the position of that arm before the movement; and so with a second or third movement. Therefore, if there were an "image" which acted as a standard it had to be one which was constantly modified by the continual postural changes. Since such an entity was scarcely an "image" in the usual almost pictorial sense of that word, Head labelled this complex standard "Schema."

Similarly, appreciation of differences between two comparable stimuli presented to two senses, cannot depend on some matching of static images against each other. Instead the writer postulates that as a result of its perceptual experiences the organism develops the capacity to react in the same way to those features of perceptual situations which may be mediated by more than one sense; and that eventually it becomes able to react to those features independently of the modality which at a given moment

happens to mediate them. That is, the organism learns to react to size, for instance, but the particular sense which subserves the perception of size at a given instant may be relatively unimportant.

This concept of Schema is in no sense the concept of an image; any of the normal uses of the word "image" imply specific sensory processes. But the whole point of introducing this amended version of the "Schema" was to obviate the need to think in terms of one or other of the senses when one considers situations where several senses co-operate. In this interpretation the Schema is a manner of functioning of the organism; it is a skill which the organism develops and which enables it to make the greatest use of its sensory equipment. The physiological processes of such a Schema are a mystery to the writer. But the full implications of the concept can be established experimentally by a host of possible investigations involving different aspects of performances in which two senses—or more—take part.

This paper has been adapted and abridged from a thesis of the same title submitted for the degree of Ph.D. at the University of London. The writer is particularly indebted to Mr. J. W. Whitfield for his constant help in supervising this experiment.

This research was assisted by a grant from the Central Research Fund of the University of London.

References

Buxton, C. E., and Crossland, H. R. (1937). The concept of "Eye-Preference." *Amer. J. Psychol.*, **49**, 458–461.

Cuff, N. B. (1931). A study of eyedness and handedness. *J. exp. Psychol.*, **14**, 164–175.

Downey, J. E. (1933). Laterality of function. *Psych. Bull.*, **30**, 109–142.

Gahagan, L. (1933). Visual dominance-acuity relationships. *J. gen. Psychol.*, **9**, 455–459.

Graham, C. H. (1951) in Stevens, S. S., *Handbook of Experimental Psychology*, Chap. XXIII, New York.

Head, H. (1920). *Studies in Neurology*, Vol. II, Oxford.

Horowitz, M. (1949). An analysis of the superiority of binocular over monocular visual acuity. *J. exp. Psychol.*, **39**, 581–596.

Jastrow, J. (1889). Perception of space by disparate senses. *Mind*, **11**, 539–554.

Woodworth, R. S. (1950). *Experimental Psychology*, London.

FIGURAL AFTER-EFFECTS, RETINAL SIZE, AND APPARENT SIZE

BY

N. S. SUTHERLAND

From the Institute of Experimental Psychology, University of Oxford.

The results of earlier experiments on the question of whether figural after-effects are affected by *apparent* as opposed to *retinal* size are shown to be inconclusive. A new hypothesis is proposed namely that both factors may be responsible for producing figural after-effects, and four experiments have been made to test it. Situations were used in which the apparent sizes of the figures were determined by the size-constancy effect. It was found that where retinal sizes of test and inspection figures are the same and apparent sizes are different, figural after-effects in the direction which would be predicted on the basis of apparent sizes are obtained. It was further shown that where retinal and apparent sizes are in conflict, whether a figural after-effect is seen or not, and the direction of the figural after-effect, depends upon the balance between these two factors.

I

Introduction

A number of experiments have been made in recent years to discover whether visual figural after-effects depend upon retinal size or apparent size. Previous experimenters seem to have assumed that the two alternatives are mutually exclusive, and they have tended to use experimental designs in which the kind of figural after-effect obtained after looking at the inspection figure would differ according to whether it was being determined by the apparent or retinal size of the figures. We wish to argue that previous experiments have been inconclusive because they have started from the assumption that figural after-effects must be either wholly dependent on retinal size or wholly dependent on apparent size, and we have performed experiments to show that this is a false dilemma and that in fact both factors are operative.

In what follows we shall use the expression *inspection figure* to refer to the figure inspection of which determines the figural after-effect. In producing a figural after-effect, a fixation point is exposed with the inspection figure and the subject fixates this point for a period of about a minute. The term *test figure* refers to the figure which is exposed to the subject after the period of fixation of the inspection figure. Its contours are made to fall near the position on the retina upon which the contours of the inspection figure fell by getting the subject to fixate a fixation point exposed with it. It is the test figure which undergoes a distortion when there is a figural after-effect. For the purposes of detecting such a distortion a third figure is often exposed at the same time as the test figure but on the other side of the fixation mark, and the subject is asked either to pass some judgment on the apparent difference between this figure and the test figure, or to adjust this figure to the same size, or position as the test figure. This is the *comparison figure*. We make the assumption throughout that in figural after-effects the contours of the test figure tend to be seen as displaced away from the position occupied by the contours of the inspection figure (*v.* Kohler and Wallach 1944). The issue at stake is how far is such displacement of the test figure governed by the retinal positions of the contours of test and inspection figures, how far by the apparent positions and sizes of the test and inspection figures.

We now proceed to a detailed consideration of previous experiments on figural after-effects in relation to apparent and retinal size. Two different methods have been employed by previous experiments on this problem:

(1) Using a method first employed by Thouless (unpublished) Prentice (1950) and Hochberg and Bitterman (1951) attempted to discover whether apparent size or retinal size was determining the figural after-effect by placing the inspection and test figures at different distances from the subject. The figures used in both experiments were circles and in both cases the more distant figure was considerably larger than the nearer. (*v.* Fig. 1 and accompanying text.) In Prentice's experiment the sizes of

(A) (B)

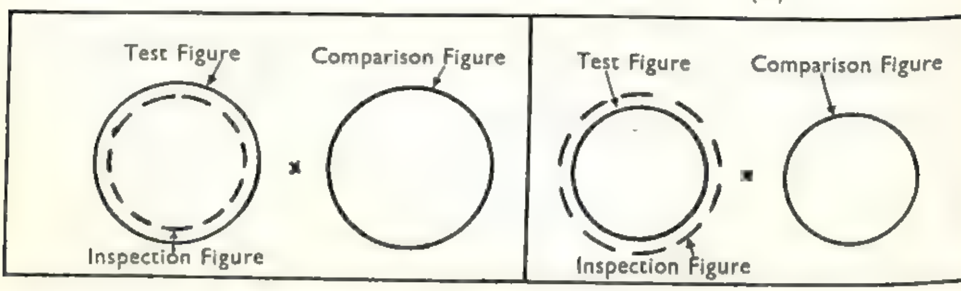


FIG. 1. Illustration of experimental design used by Prentice (1950), Hochberg and Bitterman (1951) and self (Experiment 3).

Solid outline circles represent the test and comparison figures. Broken outline circles represent the inspection figures.

In Prentice's experiment (A) depicts the relationship between *retinal sizes*—the inspection figure is retinally smaller because it is further away from the subject. (B) depicts the relationship between the *apparent sizes* of the figures—the inspection figure having larger physical dimensions than the test and comparison figures is of a larger apparent size despite its greater distance. If retinal size is the determining factor in figural after-effects then the left hand figure (test figure) should be seen as larger than the right hand figure (comparison figure). If apparent size is the determining factor then the right hand test figure should look larger than the left.

In Hochberg and Bitterman's experiment and in own experiment (3) this situation is reversed. (B) now represents relative *retinal sizes*, (A) represents relative *apparent sizes*. If retinal size determines figural after-effect the right test circle should be seen as larger than the left; if apparent size determines figural after-effect, the left test circle should be seen as larger than the right.

the figures and their distance from the subject were so arranged that while the *retinal projection* of the inspection figure fell within that of the test figure, owing to the effects of constancy the inspection figure had a larger *apparent size* than the test figure, i.e. the inspection figure was larger and further away than the test figure. In Hochberg and Bitterman's experiment this situation was reversed: the inspection figure was nearer and smaller than the test figure, and their relative sizes were such that the retinal projection of the inspection figure lay outside that of the test figure, whereas the apparent size of the inspection figure was smaller than that of the test figure. Thus in Prentice's experiment the test figure should be seen as smaller if figural after-effects depend on apparent size, as larger if they depend on retinal size, while in Hochberg and Bitterman's experiment the reverse is true. In both experiments results were obtained which unequivocally pointed to retinal size as the determining factor.

If, however, we allow that both factors may contribute to figural after-effects, in a situation in which the two factors are in conflict the result will depend presumably on

(*a*) the absolute extent to which retinal size and apparent size can contribute to figural after-effects in any situation, and (*b*) the differences in the apparent sizes and retinal sizes of the inspection and test figures in a given situation. Thus:

(*a*) It is possible that even if both retinal size and apparent size are important in determining the figural after-effect, the retinal size is the more important factor of the two and in the experiments under consideration where the two factors are in conflict any effect due to the apparent sizes of the test and inspection figures might be concealed.

(*b*) As far as can be judged, in both experiments the separation of the figures according to retinal size was about the optimal for producing figural after-effects. Since measures of constancy for individual subjects are not quoted in either of the reports of the experiments, it is impossible to say accurately what the relative apparent sizes of test and inspection figures were. Kohler and Wallach (1944) have shown that if the contours of test and inspection figure are either too close together or too widely separated figural after-effects diminish, and this may have been the case with the apparent sizes in the experiments under review. In addition to this, constancy may be diminished by prolonged fixation (though this cannot be overcome by any experimental design in which the differences in apparent size between test and inspection figures are the result of size constancy).

(2) In an earlier experiment of different design, Prentice (1947) obtained results which he interpreted at the time as favouring the hypothesis that apparent size and not retinal size was the determining factor in producing figural after-effects. His experimental design is shown in Figure 2 and the text thereto. In this experiment the inspection figure was exposed at 2 metres from the subject and the test figure was exposed for some trials at 2 metres and for others at 6 metres. If retinal size of the

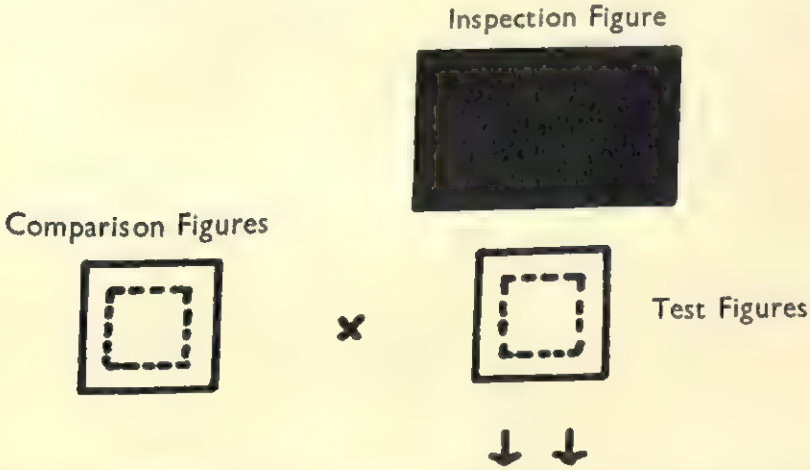


FIG. 2. Illustration of the experiment performed by Prentice (1947).

The filled in rectangle is the inspection figure. The squares with solid outlines represent retinal sizes of test and comparison figures exposed at the same distance as the inspection figure (condition 1). The squares with broken outlines represent retinal sizes of test and comparison figures when exposed further away than the inspection figure (condition 2). Arrows represent direction of displacement of test figure in figural after-effect. The issue is whether the test figure will appear displaced through the same angular distance in condition 2 as in condition 1 (retinal size operative), or through the same spatial distance i.e. a smaller angular distance (apparent size operative).

figure is the determining factor, the right-hand test square should be displaced approximately three times as much in relation to the left at a distance of 6 metres as at a distance of 2 metres. He found that in fact there was no difference in the PSE of the test and comparison squares under the two conditions. He points out, however, that this result might possibly be explained on the grounds that when the test squares are at 6 metres, the contours of the right-hand test square will, in fact, be three times as far away from the contour of the inspection rectangle as at 2 metres measured in terms of retinal size. This difference in angular separation of the test and inspection figures might lead to a decrease in figural after-effect (measured in terms of retinal displacement), and since the amount of displacement found at a distance of 6 metres was, in fact, one-third of that at a distance of 2 metres (i.e. the actual displacement of the test figure was the same in both cases), there is an alternative explanation for this result. Prentice tried to rule out this explanation by repeating the experiment with the inspection figure at a distance of 3 metres and the test figure at a distance of 3 metres, 5 metres, and 7 metres. Under these three conditions Prentice again found that the PSE for the setting of test and comparison squares remained the same. Since it seems unlikely that in all the changes of distance of the test figures made, the resulting differences in angular separation of the figures should produce exactly the same result, namely that the actual displacement of the test figure should remain the same while the angular displacement changes through varying degrees, it seems that this experiment affords some evidence for apparent size being a factor in figural after-effects. Prentice himself adopted the alternative explanation of this experiment after he had performed his 1950 experiment (described above) despite the fact that he had already attempted to rule out that explanation in the design of the experiment we are discussing. If now we treat the 1947 experiment as affording evidence for the influence of apparent size on figural after-effect and the 1950 experiment as affording evidence for the influence of retinal size, then we have no need for the alternative explanation of the 1947 experiment which Prentice himself admits to be unsatisfactory.

One important difference between the 1947 experiment and the 1950 and 1951 experiments should be noticed. In the 1947 experiment both retinal and apparent size would be working to produce a displacement in the same direction, whereas in the later experiments, in which no evidence for the importance of apparent size in figural after-effects was discovered, any effect due to apparent size would be in conflict with any effect due to retinal size.

Thus the whole issue of how far retinal size and how far apparent size influence figural after-effects can be reopened, and in reopening it we must be careful not to start from the premiss that figural after-effects are due entirely to retinal size or entirely to apparent size. Both factors may contribute, and the resulting after-effect may be due to the balance between them. If in fact perception is a two-stage or multi-stage process, then it seems likely that both factors would be of importance in determining figural after-effect, since whatever the mechanism of the figural after-effect it may be at work at each stage of the perceptual process, and the fact that figural after-effects have been shown to occur in the third dimension (Kohler & Emery 1947) indicates that figural after-effects do in fact occur at some late stage in the perceptual process.

The issue is one of considerable importance not only for arriving at the mechanism behind figural after-effects but for arriving at the brain mechanism responsible for apparent size and in particular size constancy. Thus if Kohler's explanation of figural after-effects in terms of figural currents is correct, then given his theory of apparent size, he must predict that the apparent size of figures will have some influence

on figural after-effects since the mechanism he gives for apparent size involves brain fields and a topographical projection, and within such a mechanism satiation will presumably occur. On the other hand if Osgood and Heyer's theory of figural after-effects is correct, then the discovery that figural after-effects are in part determined by apparent size will limit the kind of mechanism we can postulate to explain apparent size and the constancies. The present series of experiments was undertaken in an attempt to discover whether apparent size influences figural after-effects under optimal conditions, i.e. where no contrary effect due to retinal size would be predicted. The basic design of the experiment follows that employed by Thouless (unpublished) and Prentice (1950), but it differs from theirs in that in the crucial experiments the retinal sizes of the figures were made equal.

II

Experiments

Experiment 1. In experiments (1) and (2) an attempt was made to eliminate the influence of retinal size as much as possible by using a situation in which retinal sizes were equal, but apparent sizes were not. By this method it was hoped to isolate any effect that might be due to apparent size. In the first experiment the inspection figure was an outline circle of diameter 4 in., drawn in Indian ink, on a large white card. The test figure was an outline circle of diameter 10 in. and was exposed with a comparison figure of identical dimensions on the other side of the fixation point. Test and comparison figures were fixed to a white wall. The inspection figure was at a distance of 4 ft. 9·6in. from the subjects' eyes: the test figure at a distance of 12 ft. At these distances the angle subtended by the test and inspection figures on the retina should be the same (*v.* Fig. 3). Subjects were asked to fixate the inspection figure for a period of 45 seconds, whereupon the inspection figure was withdrawn suddenly, and subjects fixated the fixation point on the wall and were asked to say which of the two circles on the wall looked bigger to them.

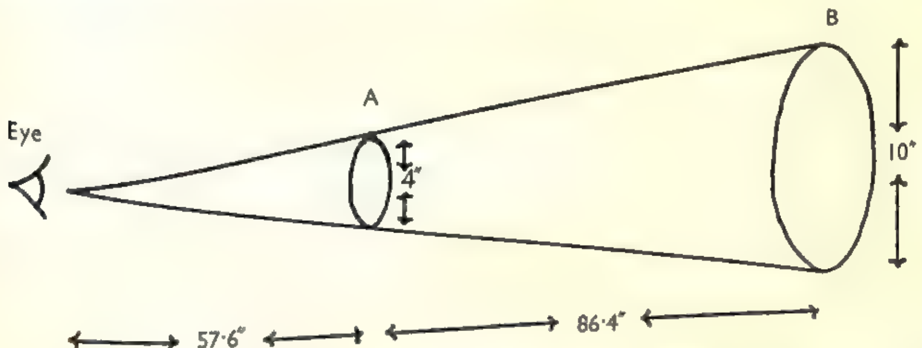


Fig. 3. The design of own experiments (1) and (2).

In experiment (1) A was the inspection figure, B the test figure. If there is a figural after-effect due to apparent size, the test figure should be seen as larger than the comparison figure, since the apparent size of A is less than B, i.e. where the inspection figure (A) is left of the fixation point the left hand of the two identical circles at (B) should be judged the bigger, where the inspection figure (A) is right of the fixation point, the right hand of the two circles at (B) should be judged the bigger. In experiment (2) B is the inspection figure, A the test figure. If there is a figural after-effect due to apparent size, the test figure at (A) should be seen as smaller than the comparison figure at (A), since the test figure (A) now has a smaller apparent size than the inspection figure (B), i.e. if the inspection figure (B) is to the left of the fixation point the right hand of the two identical circles at A should be judged bigger, if the inspection figure is to the right of the fixation point the left hand of the two circles at A should be judged bigger.

All instructions were given before the experiment began. With half the subjects the inspection figure was to the left of the fixation point, with the other half to the right. The distance between the fixation point and the edge of the circle was 0·75 in. on the inspection figure and 2·25 in. on the test figure. Each subject gave only one judgment.

Experiment 2. In the second experiment the inspection figure was the far figure, and the test figure was the near one, i.e. one circle only was first exposed at a distance of 12 ft. to right or left of the fixation point, and two figures (test and comparison) were then shown at 4 ft. 9·6 in. Subjects were asked to close their eyes for a period of just over a second on the termination of the 45 second inspection period in order to prevent them seeing the test figures before they were in position. Otherwise the procedure was exactly the same as for experiment 1.

Experiment 3. The third experiment was undertaken in order to confirm the results of Prentice (1950) and Hochberg and Bitterman (1951) i.e., in order to discover what happens when retinal size and apparent size are in conflict. In this experiment the inspection figure was the near figure as in experiment 1, but it was exposed at a distance of 4 ft. (*v.* Fig. 1).

Experiment 4. In the fourth experiment an attempt was made to determine for individual subjects the distance from the eye at which the test figures appeared equal after inspection of the figure at 12 ft. Each subject made 70 judgments on the relative apparent sizes of the two circles exposed on the test card. The distance of the test card from the subjects' eyes was varied, 7 different distances being used: 10 presentations were made at each distance in a random order. The order was varied from subject to subject. The procedure was as follows: after a preliminary inspection period of 45 seconds, the subject closed his eyes for a period of just over a second while the test figure was placed in position. On the command "Look" given by the experimenter, subjects opened their eyes fixated the fixation point on the test card and gave their judgment "Left bigger," "Right bigger," or "Equal" as quickly as possible. As soon as the judgment had been made, the test card was removed, subjects fixated the inspection figure for a period of 10 seconds and the procedure was repeated. Unfortunately, all judgments had to be made in one session. After successive groups of 14 judgments, subjects were given 10 minute rest periods. Two subjects had to be discarded because their eyes grew tired with the successive fixations before the experiment had been concluded. The experiment was conducted with the inspection circle on the left of the fixation point with 3 subjects, and on the right with the other 3.

Care was taken in exposing the figures not to introduce contours irrelevant to the experiment anywhere near the parts of the visual field upon which the figures were projected. All experiments were carried out in full daylight, though no direct sunlight was allowed to enter the room.

All the experiments were performed with binocular vision. Subjects with defective sight wore glasses to correct for it. None of the subjects used in experiment 4 had defective vision.

Subjects for the experiments were all volunteers. They were of both sexes and ages ranged from 18–32. All subjects were experimentally naïve. Some of the subjects were students, nurses or secretaries; many were foreign visitors to Oxford.

III

Results

The results of the first three experiments are presented in Table I.

It will be seen from this Table that altogether 5 subjects failed to comply with the instructions given them and refused to give a judgment to the effect that one circle looked bigger than the other. For this reason these subjects have not been included in the statistical analysis. It is sufficient for our purposes to show that where subjects do see a difference in the size of the test and comparison circles, the direction of that difference is correlated with the side upon which they have previously viewed the inspection figure. A Chi squared test was performed on each of these experiments

to discover whether the difference in the distribution of judgments where the inspection figure was on the left and where it was on the right was significant. The results of experiments (1) and (3) were significant at well beyond the 0·01 level of confidence, and those of experiment (2) were significant at the 0·001 level of confidence. The results of experiments (1) and (2) were in the direction which would be predicted if apparent size was a determining factor in this situation in producing the figural after-effect. In experiment (1), the test circle is seen as larger than the comparison circle presumably because in this experiment the inspection circle was of smaller apparent size than the test circle. In experiment (2), the test circle is seen as smaller than the comparison circle presumably because in this experiment the apparent size of the inspection circle is greater than that of the test circle. In both cases therefore the contours of the test circle moved away from the position of the contours of the inspection circle, if we consider the apparent sizes of the circles and not their retinal sizes which were equal.

TABLE I

THE RESULTS OF EXPERIMENTS 1, 2, AND 3

Experiment	Position of Inspection Figure	Number of Subjects	*Judgments on Test Figures* Left Bigger	Right Bigger	X^2	Level of Significance	Subjects judging Test Figures as equal
1	Right	12	3	9	8·4	0·01	2
	Left	12	10	2			
2	Right	11	10	1	11·7	0·001	1
	Left	11	2	9			
3	Right	10	8	2	9·9	0·01	2
	Left	10	1	9			

The results of experiment (3) in which a similar situation was employed to that used by Hochberg and Bitterman (1951) are those that would be predicted on the basis of retinal size. As in Hochberg and Bitterman's experiment, retinal size and apparent size considered as factors in producing a figural after-effect are in conflict in this experiment, since moving the inspection figure nearer to the subject makes it retinally larger than the test figure though its apparent size remains smaller than that of the test figure. Thus the result of this experiment confirms the results already obtained by Hochberg and Bitterman and by Prentice viz. that in some situations retinal size may predominate over apparent size in producing a figural after-effect.

Subjects were asked at the end of the experiment to describe any changes that occurred in the inspection figure during the period of inspection. Most subjects reported changes of the type reported by Kohler and Wallach in their original monograph (1944). Changes reported (in order of frequency) were: (*a*) fading of different parts of the circle; (*b*) distortions of the circle's shape, e.g. the circle looked polygonal at times to some subjects, to others it appeared flattened out either in the vertical or horizontal directions; (*c*) a few subjects reported that the outlines of the circle

appeared to be moving in a circular direction; and (*d*) one subject reported that the circle appeared to be moving in the third dimension, i.e. to be alternately standing out from the wall to which it was attached and moving back to the wall. It is perhaps worth noting that a smaller proportion of the subjects who gave judgments at odds with the majority or who were unable to give a judgment at all saw such changes than of those subjects who gave judgments which tallied with the majority's judgments.

(4) Results for the fourth experiment are given in Table II.

TABLE II

THE RESULTS OF EXPERIMENT 4

	Inspection Circle on Left									*Inspection Circle on Right*								
Subject	1			2			3			4			5			6		
Distance from eyes in inches	L	=	R	L	=	R	L	=	R	L	=	R	L	=	R	L	=	R
47·6	4	0	6	10	0	0	2	2	6	2	6	2	2	4	4	2	6	2
51·6	0	0	10	1	3	6	3	3	4	4	6	0	4	2	4	4	6	0
55·6	0	0	10	0	8	2	2	5	3	6	4	0	10	0	0	6	4	0
57·6	0	0	10	0	2	8	0	3	7	8	2	0	10	0	0	8	2	0
59·6	0	0	10	0	2	8	0	1	9	10	0	0	9	1	0	7	3	0
63·6	0	3	7	1	0	9	1	2	7	7	3	0	10	0	0	10	0	0
67·6	0	1	9	0	0	10	0	1	9	10	0	0	10	0	0	10	0	0

The results of experiment (4) in general confirm our hypothesis. The basic situation is the same as in experiment (2). A figural after-effect which depended solely on apparent size would result in subjects seeing the test circle as smaller than the comparison circle since owing to constancy the apparent size of the inspection circle was larger than that of the test circle at all positions of the test circle. If retinal size alone were the determining factor the test circle should be seen as larger than the comparison circle at distances from the eye of less than 4 ft. 9·6 in., as smaller at distances beyond this. It is obvious that for all 6 subjects who completed the experiment the point at which the test figure has to be exposed in order that the comparison and test circles should appear equal is considerably nearer the eye than the point at which inspection and test figures are retinally equal (4 ft. 9·6 in.). The results are disappointing, however, for two reasons: (1) There are too many inconsistencies in the results of individual subjects for it to be worth attempting to determine the point for each subject at which the probabilities of obtaining a judgment of "left bigger" or "right bigger" are equal. (2) With several of the subjects the proportion of judgments which were in the opposite direction to the judgments which we would expect if apparent size were the determining factor became less as the experiment proceeded. This was particularly the case with subjects (1) and (6). It may be that subjects tended to repeat the judgment which they found to be most frequent, and owing to the design of the experiment this would in fact be the judgment based on apparent size. It should be noticed that none of the subjects realized that their judgments were being influenced by the position at which the test figure was exposed, though some of the subjects realized towards the end of the experiment that it was in fact being exposed at different distances from the eye.

IV

Discussion

The results of experiments (1) and (2) unambiguously support the hypothesis they were designed to test viz. that it is not merely retinal size that is important in determining figural after-effect but that the relative apparent sizes of test and inspection figures may also be a determining factor. The situation is, however, complicated by two further considerations:

Firstly, Kohler and Wallach found that when inspection and test figures were the same in size and when they were exposed at the same distance from the eye, either the test figure was seen as very slightly smaller than the comparison figure or the test and comparison figures were seen as of the same size. Now in experiment (2) the result might possibly be explained without introducing the factor of apparent size as part of the explanation since here the test figure is seen as smaller after inspection of the inspection figure whose contours fall on the same part of the retina as those of the test figure. The degree to which subject saw the test figure as smaller than the comparison figure in experiment (2) indicates that this is not, in fact, the correct explanation, since Kohler and Wallach found only a very small effect in this situation where test and inspection figures were in the same plane. Moreover, experiment (1) cannot be explained in this way since in this experiment the test figure is seen as larger than the comparison figure.

Secondly, we have assumed so far that figures of the same shape and angular size will, in fact, form identical images on the retina irrespective of the distances from the eye at which they are exposed. This is not the case. Duke-Elder (1932) states that as accommodation increases the size of the retinal image for a figure of given angular size increases. This fact, cannot, however, be used to explain our findings. Thus: (*a*) At the distances we are concerned with the effect is minute. (*b*) Any effect due to this factor would operate in the opposite direction to that due to apparent size, since while the retinal size of the near circle will be increased relatively to that of the further through the effects of accommodation, the effect that was demonstrated in experiments (1) and (2) was in the direction which would be predicted if the further circle is seen as larger than the nearer.

Experiment (3) is in effect a repeat of the experiment performed by Hochberg and Bitterman (1951). Our results show that with the degree of retinal separation of test and inspection figures produced by exposing them at a distance of 4 ft. and 12 ft. respectively, retinal size and not apparent size is the determining factor in producing the after-effect. This raises the question of what happens when the smaller figure is at intermediate distances. If both effects (i.e. that due to retinal size and that due to apparent size) are interacting we would predict that as the difference in retinal size is increased by moving the position of the smaller circle nearer to the eye and away from the point at which retinal size is approximately the same (where the angular distance of test and inspection figure is the same), there should be a decrease in judgments of the kind that would be expected on the basis of apparent size and an increase of the judgments of the kind that would be expected on the basis of retinal size. Moreover, since constancy is known to vary considerably from subject to subject we would predict that there might be considerable differences in the judgments given by different subjects with the smaller circle at varying distances, and in particular if both effects are occurring simultaneously we would expect that there would be a point for each subject where the two effects exactly counter-balance one another, and there is no figural after-effect, i.e. test and comparison figures are judged to be the same size.

The results of experiment (4) afford some confirmation for this prediction, but for the reasons given in the last section these results are not completely satisfactory. This experiment should be repeated using the method of adjustment instead of that of constant stimuli, since that would prevent subjects becoming fixated on a given response and it might be possible to determine with some accuracy for different subjects what distance the test figures would have to be exposed at in order that they should be seen as equal after the inspection periods.

As has already been indicated in the Introduction, the hypothesis that we regard as having been proved by the experiments we have undertaken enables us to explain the apparently conflicting results of earlier experiments and these experiments may be regarded as affording further confirmation of the hypothesis, namely that both retinal size and apparent size are of importance in determining the direction and amount of a given figural after-effect.

In summary then, experiments (1) and (2) prove that apparent size is one determining factor in producing figural after-effects. Experiments (3) and (4) indicate that retinal size is also of importance in determining figural after-effects and that the two factors may both be operative at the same time in producing a given after-effect.

The implications of this discovery are:—Firstly, perception is at least a two stage process otherwise it is impossible to explain how both factors can contribute to the figural after-effect at one and the same time. Secondly, whatever the mechanism behind the perception of apparent size and figural after-effects, the two mechanisms must be such that at some stage in the perception of apparent size the mechanism behind the production of figural after-effects can operate. It is hoped to present a second paper examining in detail the theoretical implications of the discovery that figural after-effects are influenced by apparent size, both for the mechanism behind figural after-effects and for that behind the constancies.

This research was made possible by a grant given me by the Institute of Experimental Psychology. My thanks are due to Professor G. Humphrey for affording me the facilities necessary for the conduct of the experiment, and to Mr. J. A. Deutsch for much helpful advice.

References

Duke-Elder, W. S. (1932). *Text Book of Opthalmology*. London. Vol. 1, p. 745.

Hochberg, J. E., and Bitterman, M. E. (1951). Figural after-effects as a function of the retinal size of the inspection figure. *Amer. J. Psychol.*, **64**, 99–102.

Kohler, W., and Emery, D. A. (1947). Figural after-effects in the third dimension. *Amer. J. Psychol.*, **60**, 159–201.

Kohler, W., and Wallach, H. (1944). Figural after-effects. *Proc. Amer. Phil. Soc.*, **88**, 4, 269–357.

Osgood, C. E., and Heyer, A. W. Jr. (1952). A new interpretation of figural after-effects. *Psychol. Rev.*, **59**, 98–118.

Prentice, W. C. H. (1947). The relation of distance to the apparent size of figural after-effects. *Amer. J. Psychol.*, **60**, 617–23.

Prentice, W. C. H. (1950). On the size of the figural after-effect with varying distances. *Amer. J. Psychol.*, **63**, 589–93.

Proceedings of the Experimental Psychology Group, 1953

13TH–14TH APRIL, 1953.—Extended Meeting at Nottingham. *1st Session:* "The Origin of Adaptive Behaviour," by W. Ross Ashby (by invitation). Short communications by members of the Psychology Laboratory, University of Nottingham—L. J. Newson, Elizabeth Newson and G. Collins (all by invitation). *2nd Session:* "The Perceiving and Remembering of Instructional Television Programmes," by M. D. Vernon. "An Investigation of Intersensory Facilitation," by J. R. Symons (by invitation). *3rd Session:* "The Waning of an Instinctive Response," by R. A. Hinde (by invitation).

7TH–8TH JULY, 1953.—Extended Meeting at Cambridge. *1st Session:* "Some Maze Learning Experiments with Human Subjects," by J. von Wright. *2nd Session:* "Factors Affecting Immediate Memory," by J. Brown (by invitation). "The Use of Statistical Models in Psychology," by N. L. Webb (by invitation). *3rd Session:* Symposium on "Psychological Theories and Animal Behaviour," by H. Verplanck, J. S. Kennedy (both by invitation) and S. Koch. *4th Session:* Symposium (continued). Discussion opened by D. Russell Davis.

18TH–19TH SEPTEMBER, 1953.—Extended Meeting at Reading. *1st Session:* "Subception and Automisation in Relation to Theories of Inhibition," by D. E. Broadbent. Demonstration by R. C. Oldfield. *2nd Session:* "Experimenting with Instructional Films," by S. Laner (by invitation). "Subliminal Auditory Perception," by N. F. Dixon (by invitation). *3rd Session:* "Dogmatism and Rigidity as Determinants of Cognition," by Milton Rokeach (by invitation).

19TH DECEMBER, 1953.—7th Annual General Meeting at University College, London. *1st Session:* "Memory Systems in *Octopus vulgaris* Lamarck," by B. B. Boycott (by invitation). *2nd Session:* "A New Type of Behaviour Theory," by J. A. Deutsch (by invitation).

Members

BABINGTON SMITH, B., O.B.E., M.A., Institute of Experimental Psychology, 34, Banbury Road, Oxford.
BARLOW, H., M.D., Department of Physiology, Downing Street, Cambridge.
BROADBENT, D. E., M.A., Medical Research Council Unit for Applied Psychology, 15, Chaucer Road, Cambridge.
CANE, Miss V. R., M.A., Medical Research Council Unit for Applied Psychology, 15, Chaucer Road, Cambridge.
CARPENTER, A., M.B., Medical Research Council Unit for Applied Psychology, 15, Chaucer Road, Cambridge.
CHAMBERS, E. G., M.A., Psychological Laboratory, Downing Street, Cambridge.
CRAWFORD, Miss A., Ph.D., Department of Psychology, The University, Liverpool, 7.
CROWN, S., Ph.D., 81, Elgin Avenue, London, W.9.
DAVIS, D. RUSSELL, M.D. (*Editor*), Psychological Laboratory, Downing Street, Cambridge.
DEUTSCH, J. A., M.A., Institute of Experimental Psychology, 34, Banbury Road, Oxford.
DREW, Prof. G. C., M.A., Department of Psychology, The University, Bristol.
FOSS, B. M., M.A., Department of Psychology, Birkbeck College, Malet Street, London, W.C.1.
GREGORY, R. L., B.A., Psychological Laboratory, Downing Street, Cambridge.
GRINDLEY, G. C., M.A., Psychological Laboratory, Downing Street, Cambridge.
HALL, K. R. L., D.Phil., Barrow Hospital, Barrow Gurney, nr. Bristol.
HARPER, R., Ph.D., Department of Psychology, The University, Leeds, 2.
HARRIS, A. D., M.R.C.S., (*Hon. Secretary*), Psychological Laboratory, Downing Street, Cambridge.
HEIM, Miss A. W., Ph.D., Psychological Laboratory Downing Street, Cambridge.
HICK, W. E., M.D. (*Hon. Treasurer*), Psychological Laboratory, Downing Street, Cambridge.
HIMMELWEIT, Mrs. H. T., Ph.D., London School of Economics, Houghton Street, Aldwych, London, W.C.2.
HINDE, R. A., D.Phil., Ornithological Field Station, Madingley, Cambridge.
HOTOPH, N., M.A., London School of Economics, Houghton Street, Aldwych, London, W.C.2.

JAMES, P. H. R., M.A., Department of Psychology, University College, Gower Street, London, W.C.1.

JONCKHEERE, A. R., B.Sc. (*Committee*), Department of Psychology, University College, Gower Street, London, W.C.1.

JOYNSON, R. B., M.A., (*Committee*), Department of Philosophy, The University, Nottingham.

KAY, H., Ph.D. (*Committee*), Institute of Experimental Psychology, 34, Banbury Road, Oxford.

LEONARD, J. A., Ph.D., Medical Research Council Unit for Applied Psychology, 15, Chaucer Road, Cambridge.

MACKWORTH, N. H., M.B., Ph.D., Medical Research Council Unit for Applied Psychology, 15, Chaucer Road, Cambridge.

MACNEILL, Miss F. E., Ph.D., Department of Psychology, The University, Manchester.

OLDFIELD, Prof. R. C., M.A., Department of Psychology, The University, Reading.

PICKFORD, R. W., D.Litt., Ph.D., Department of Psychology, The University, Glasgow, W.2.

PIERCY, M. F., M.A., Department of Experimental Psychiatry, The University, Birmingham.

POULTON, E. C., M.B., Medical Research Council Unit for Applied Psychology, 15, Chaucer Road, Cambridge.

RUSSELL, Prof. R. W., Ph.D., Department of Psychology, University College, Gower Street, London, W.C.1.

SEMEONOFF, B., Ph.D., Department of Psychology, The University, Edinburgh.

SUMMERFIELD, A., B.Sc., Department of Psychology, University College, Gower Street, London, W.C.1.

SZAFRAN, J., Ph.D., Psychological Laboratory, Downing Street, Cambridge.

TRIST, E., O.B.E., M.A., Institute of Human Relations, 2, Beaumont Street, London, W.1.

VERNON, Miss M. D., D.Sc., Department of Psychology, The University, Reading.

VINCE, Miss M. A., M.A., Psychological Laboratory, Downing Street, Cambridge.

WELFORD, A. T., M.A., Psychological Laboratory, Downing Street, Cambridge.

WHITFIELD, J. W., M.A. (*President*), Department of Psychology, University College, Gower Street, London, W.C.1.

ZANGWILL, Prof. O. L., M.A., Psychological Laboratory, Downing Street, Cambridge.

Visiting Foreign Members

NEFF, Prof. W. D., Ph.D., Office of Naval Research, Keysign House, 429, Oxford Street, London, W.1.

VON WRIGHT, J., M.A., Institute of Experimental Psychology, 34, Banbury Road, Oxford.

THE QUARTERLY JOURNAL OF
EXPERIMENTAL PSYCHOLOGY

Vol. VI 1954 Part 2

STRENGTH DURATION CURVES FOR ELECTRICAL STIMULATION OF THE HUMAN EYE

BY

C. I. HOWARTH

(From the Institute of Experimental Psychology, University of Oxford)

Baumgardt's surprising report of a minimum at about 30 milliseconds in the strength-duration curve for electrical stimulation of the human eye is confirmed. Arguments are presented against his explanation of the phenomenon.

An alternative hypothesis is developed: that the complex strength-duration curve is due to the interaction of on and off processes, which are separately excited by the make and break of the current: that these processes summate when separated by 30 milliseconds but are mutually inhibitory at 70 milliseconds: that these processes are the same as those known to be aroused by light. Two new experiments are designed to test this hypothesis, and the predicted result has been obtained from each.

Experiments are devised which show that the site of stimulation is the retina.

The accuracy of the method is discussed. Over the range 10 to 100 milliseconds distortion is found to be small. For shorter pulses distortion is to be expected in all cases, and it is suggested that the chronaxie is an expression merely of the condenser effect of the tissues. The use of this term should be discontinued in connection with stimulation through the skin.

Some parallels between these results and brightness comparison experiments with light are pointed out. The nature and location of the interaction is discussed.

I

Introduction

A new method of approach to many visual problems has recently been developed by Motokawa and his school (Motokawa, 1949, a, b & c, 1950, 1951; Gebhard, 1953). They have used the electrical threshold of phosphenes as an index of the excitability of the visual system and have discovered that this excitability is raised over a period of seconds after exposing the eye to a flash of light. This increase of excitability is related to colour processes. The electrical threshold is lowest at 1 second after a flash of red light, 2 seconds after a flash of green light, and 3 seconds after a flash of blue light (Motokawa, 1949a). The threshold is lowered still more when colour or brightness contrast can be expected. The method has been used to investigate two main problems: colour vision—particularly the old problem of the classes of colour receptors; and shape perception, by using the contrast phenomena. Motokawa claims that the method is capable of showing the part played by the retina in both these fields and supports this by experiments with frogs' and cats' eyes. (Motokawa, Iwama & Tukahara, 1950; Motokawa, Iwama & Ebe, 1952).

Unfortunately this work is almost completely unsupported so far by other workers. But this is not the only difficulty. When one reads the older literature on the electric phosphenes one finds that there has been almost no agreement on the simple electrophysiological properties. For example, there is a controversy about the effect of dark adaptation on the electrical threshold which has still not

been settled (Bouman, 1935; Achelis & Merkulow, 1930; Schwarz, 1947; Barlow, Kohn & Walsh, 1947; Abe, 1950); the chronaxie values quoted range from 0·5 to 20 milliseconds (Achelis & Merkulow, 1930; Bourguignon & Déjean, 1926; Verrijp, 1924; Iwama, 1949); and strength-frequency curves of almost infinite variety have been obtained (Rohracher, 1935; Schwarz, 1947; Abe, 1950; Barnett & Posner, 1941; Pollock & Meyer, 1938). Motokawa's school has investigated the chronaxie of the eye (Iwama, 1949) and the strength-frequency curve (Motokawa & Iwama, 1950; Abe, 1950; Motokawa & Ebe, 1952), but have suggested no explanation for the previous disagreements. In fact Motokawa is now in a minority of one about the shape of the strength-frequency curves. The number of unaccountable anomalies seems to be increasing rather than decreasing at the moment.

In addition to this there must be many doubts about the accuracy of a method of stimulation with electrodes so far away from the excitable tissue. A great deal of the imposed current must be missing the nervous structures completely. Any change in the amount of this wastage will of course appear as a raising or lowering of the threshold. There has been no systematic approach to this problem.

The site of stimulation is unknown. Motokawa considers that it is somewhere in the retina because he can repeat some of his results with excised eyes. Iwama (1949) considers that it is the ganglion cells which are excited because of the *long* chronaxie which he finds. Copolongo (1948) on the other hand thinks that central structures, even the cortex, are involved because, he claims, blind subjects experience the phosphene and have the same *short* chronaxie as normal subjects. Baumgardt (1951b) on anatomical grounds considers that it is the optic nerve which is stimulated close to its entry into the eye.

Recently Baumgardt (1951a) has shown that the strength-duration curve for electrical stimulation of the eye has more than normal interest. This work, although less well known than that of Motokawa seems to be just as important. Baumgardt claims that the method shows up very clearly the existence of a slow inhibitory process in the visual system. Because of this the strength-duration curve is complex with the threshold for long pulses being, surprisingly, higher than that for shorter pulses. This complexity has been found by no other worker, but in view of the apparent relationship between contrast and inhibition it was thought that there might be a relationship between these findings of Baumgardt and those of Motokawa.

On these grounds it was thought useful to investigate again the simple strength-duration curve for the human eye. The immediate aim of the work was to discover the nature and cause of the complexity in this curve and to test Baumgardt's hypothesis about it. It was hoped at the same time to find the site of stimulation, test the accuracy of the method and clear up some of the previous discrepancies.

II

Apparatus

Except for the single experiment reported using constant voltage, all the stimulation was done by means of constant current generators. The arguments in favour of this method of stimulation are given later in the text. Most of the results have been obtained from a specially constructed output unit giving a constant direct current from 1–10,000 microamps provided the voltage developed is less than 200 volts. When constant voltage measurements were required a 400 ohm resistance was placed in parallel with the subject whose resistance is about 5000 ohms. Attached to the stimulator was a waveform generator which gave most of the pulses needed. To obtain two pulses in opposite directions a Lucas pendulum contact breaker was incorporated in the apparatus. A 300 volt source with a series resistance of 1½ megohms gave a constant current pulse of

7 milliseconds duration. At a given interval before or after this pulse the main stimulus generator could be triggered and gave a 10 millisecond pulse of accurately adjusted intensity.

Ag/AgCl and Zn/$ZnSO_4$ electrodes were used at first but such precautions were found unnecessary when using constant current pulses. Large aluminium electrodes padded with saline soaked cotton wool were used normally. A variety of positions was tried including wick and contact lens electrodes directly on the eye. No differences could be detected in the results from the various methods and generally one large electrode on the forehead and another on the hand was used. Zn,$ZnSO_4$ electrodes were used for the constant voltage stimulation.

A Cossor double beam oscilloscope was used to check the waveforms produced and to estimate the capacity of the tissue condensers.

Method

The method used was essentially that developed by Motokawa and very fully described in his papers. The basic measurement is the phosphene threshold. The phosphene is easily seen by nearly all subjects at about 300 microamps. The phosphene is an unfamiliar experience, and it is always necessary to give the subjects a good look at it before they proceed to the threshold determination. The procedure adopted for this is always the same. The strength of the current is reduced in steps of 10 microamperes. At each current strength the subject is warned about half a second before receiving the stimulus. Usually the experimenter says "coming" and then presses a button which triggers the stimulator. More mechanical and impersonal procedures have been tried; they seem to be without effect on the results and are in general too cumbersome and inflexible. On each presentation of the stimulus the subject reports "yes" or "no" depending on whether or not the phosphene is seen. If the answer is "yes" the strength of the current is reduced one step. If the answer is "no" the stimulus is repeated at the same intensity. If more than half the answers at this intensity are "yes" the intensity is again lowered. If 3 successive "nos" occur the series is terminated and the intensity one step above is taken as the threshold. In general subjects make up their minds quite definitely when the stimulus is no longer visible, and very little time is wasted saying alternately "yes" and "no." When this does occur it is usually because the subject is trying too hard to see the last trace of the phosphene and is unable to distinguish it from the entoptic illumination. This is much more likely to occur in the dark. None of the author's subjects could approach the accuracy reported by Motokawa for the dark adapted eye. For the light adapted eye the task is an easy one and a very satisfactory degree of accuracy is attained by most subjects.

A set of stimuli always starts well above the threshold. The actual amount above the threshold is varied in a random manner. The speed at which the stimuli are presented is adjusted to give the subject the least strain. Usually a rhythm of presentation develops with the stimuli following each other at intervals of 2 to 10 seconds. Blinking interferes considerably with the phosphene but the rhythmic presentation, together with the half second warning before each stimulus, allows the subject to inhibit his blinking at the moment when the stimulus is given. Motokawa has shown that the speed at which the stimuli are presented affects the threshold. But the amount of this effect is small compared with the size of the phenomena investigated and apart from the regular presentation of the stimuli no attempt has been made to control this factor. The threshold is always approached from above. Determinations using increasing current strengths have been tried. In general a slightly higher threshold results but the task is much more difficult under these conditions.

At the beginning of the investigation a diffusing screen illuminated from behind was used and the effect of different illuminations investigated. No change in the threshold could be found over a very wide range from bright sunlight to near the photopic threshold. Below the photopic threshold a big change occurred which is described later. Nothing else of interest was found so the investigation is not reported in full. It is mentioned here as the justification for the absence of accurate control of illumination in the later experiments. The general illumination of the room never fell below 3 millilamberts and the subjects were asked to fixate a spot on the whitewashed laboratory wall. This was by no means uniformly illuminated but accuracy was better and fatigue less under these conditions than when using a diffusing screen.

Subjects vary considerably in their ability to repeat the threshold determinations.

TABLE I

CONDITION A

Subject	Time "t" milliseconds											
	10	15	20	25	30	40	50	60	70	80	90	100
CIH	165	125	112	97	100	118	140	165	175	—	175	—
DB	175	150	120	102	100	107	120	160	173	183	190	190
H.B.G. ..	130	120	108	105	100	110	125	145	155	160	155	160
Average ..	157	132	113	102	100	112	128	157	167	171	173	175

TABLE II

CONDITION B

Subject	Time "t" milliseconds											
	10	15	20	25	30	40	50	60	70	80	90	100
CIH	136	114	104	97	100	119	147	149	150	150	130	138
DB	—	115	107	—	100	107	132	132	132	—	136	136
P	133	117	105	102	100	109	129	135	126	—	120	129
Average ..	135	115	106	99	100	112	136	139	136	—	129	135

TABLE III

CONDITION C

Subject	Time "t" milliseconds									
	10	20	30	40	50	60	70	80	90	100
CIH	159	131	100	90	95	107	120	124	113	107
DB	177	122	100	100	108	126	137	130	130	130
JP	131	105	100	100	107	121	131	128	138	140
Average ..	155	119	100	97	103	118	129	127	127	126

Tables I, II and III.

These are the results from which curves A, B and C of Figure I are plotted. Each figure represents the average of at least two determinations. To enable the results from different subjects to be pooled the thresholds are reduced to a percentage of the threshold for "t" = 30 milliseconds.

Usually however the scatter obtained is less than 20 per cent. and in most cases a very much higher accuracy can be attained after very little practice.

As the session proceeds there is occasionally a slight fall in the threshold values after about 5 minutes. Very commonly if the subjects are kept longer than ¾-hour a sudden sharp rise in the threshold occurs. Once this has started it tends to continue and these readings are always discarded. To check that neither of these phenomena is affecting the shape of the curve, the mid-point (usually at 30 milliseconds) is checked repeatedly during the course of the experiment. If any change occurs the intermediate readings are always discarded. As a further check against slight drift of the threshold the points on the curve are always determined in series in both directions. In general there is remarkably little drift during the usual experimental session of half an hour.

The threshold tends to vary somewhat from day to day in the same subject, over a range sometimes as great as 50 per cent. but usually the readings are stable within about 5 per cent. provided the placing of the electrodes is unchanged. When results from different days are pooled allowance is always made for this effect. In the results reported the thresholds are given as a percentage of the threshold for the 30-millisecond pulse.

Many subjects have been used. None of them have given results contrary to those predicted by the author's hypothesis. Two out of about 30 have found the task extremely difficult and were unable to give sufficiently repeatable results for any curve to be obtained. This was in one case due to a sleepless night, in the other no adequate reason could be found. The results reported are not selected because they are better than most, but because they are from subjects who were prepared to come several times, and who have been tested over the full range of the curve reported.

III

Experiments

Strength-duration curve

The strength-duration curve for the human eye was done repeatedly and carefully using constant current square pulses. (Table I, Figure 1 curve A). The results amply confirm those of Baumgardt (1951a). There was always an increase in the threshold current for pulses longer than 30 milliseconds. The amount of this rise varied from person to person, in some cases being as much as 75 per cent. and in others only 30 per cent. The threshold frequently fell again for pulses longer than 70 milliseconds, but this part of the curve has not been accurately explored.

Baumgardt's Hypothesis

Baumgardt has pointed out that the raising of the threshold for pulses longer than 30 milliseconds can only be due to an active process, presumably in the nerve tissue. Such an effect could not be produced by any amount of distortion of the pulse by the tissues surrounding the eye (discussed on page 56). He has postulated (1951a) that a slowly developing inhibitory process is responsible for the rise in threshold. This hypothesis had already been very fully developed by Baumgardt and Segal (1947) in connection with some experiments with light.

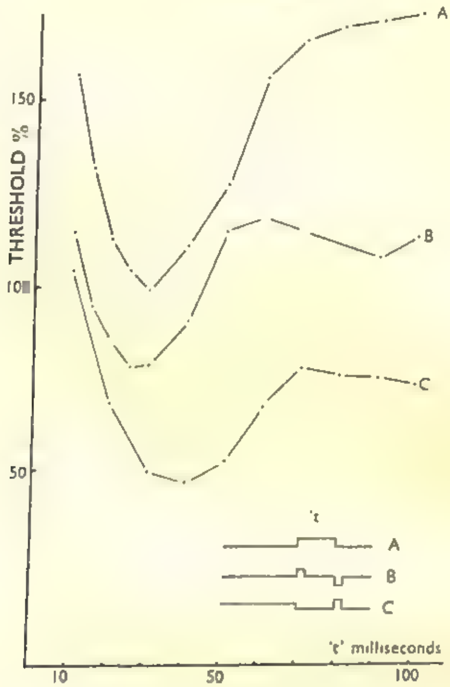
The following is an extract from their summary of the paper. "The stimulus arouses in an excitable element or system two antagonistic processes, one facilitatory and developing rapidly, the other inhibitory, released by the first and developing later and probably more slowly. Because of this difference in the time constants the sensation passes through two successive stages. During the first the behaviour of the nervous cell is primarily governed by the growth of the facilitatory function; during the second the facilitation has reached a maximum while the inhibitory function is still growing. An extra influx of energy increases most the developing

process. As a result the initial supernormal phase is followed by a subnormal phase."

The inhibitory process is "retroactive" so that the strength of the sensation does not depend on the maximum level of excitation reached by the system but is an average of the total excitation.

This hypothesis seems to have difficulty in accounting in the same terms for the strength-duration curves for single and for double pulse stimulation. Baumgardt (1951b) used two short pulses of constant duration with a varied interval between them. He found the threshold, and therefore the inhibition, to be a

FIGURE 1.



Threshold as a function of time "t" for the various wave forms shown.

The threshold is given as a percentage of the threshold for t = 30 milliseconds. Curve B is displaced downwards 20 per cent. and curve C 50 per cent.

Each point represents the average for the three subjects in Tables I, II and III.

maximum at 30 milliseconds. The present author has completely confirmed Baumgardt's experimental results (4 subjects Figure 2). However when the theory is applied to the strength-duration curve for a single pulse, inhibition is expected to be at a maximum when the upward slope of the curve is greatest, i.e. at about 45 milliseconds. One would like to know whether this is a real difference which needs to be explained or whether it is the result of a fault in the theory.

This theory is also weak because of the number of specific but untested assumptions which it makes. It is hard to think what kind of experiments would be capable of demonstrating the slowly developing inhibitory process which is also retroactive. An alternative hypothesis has been developed which is simpler, more in accord with known physiological processes and seems to have greater predictive and explanatory power. It is also more easily testable by experiment.

A New Hypothesis

The new hypothesis explains the complexity of the curve as follows:—

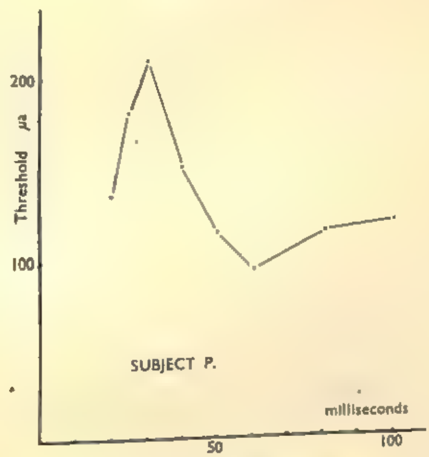
(i) The single square electrical pulse is being converted by adaptation into two separate pulses of opposite polarity, one occurring when the current is switched on and the other when it is switched off.

(ii) These pulses separately excite two processes called "on" and "off" processes.

(iii) These processes later interact by summation or inhibition, the nature of the interaction depending on the time interval between them. At 30 milliseconds they summate maximally whereas at 70 milliseconds they are mutually inhibitory.

(iv) These "on" and "off" processes are the same as those aroused by light.

FIGURE 2.



Threshold as a function of the time separating two 10 millisecond pulses. Similar results have been obtained from 3 other subjects.

This hypothesis was suggested in the first place by the symmetry between the strength-duration curve for a single square pulse (Figure 1 A) and the strength-duration curve for two 10 millisecond pulses with a varied interval between them. (Figure 2). The first has a minimum at 30 milliseconds and the second a maximum at the same interval. It suggested that there was a great similarity between the two conditions of stimulations from the point of view of the nervous structures involved. This could only be the case if the continuous stimulation of the single prolonged pulse were converted into periodic stimulation by the adaptation of the tissues. Such adaptation, as is well known, will be followed by a rebound response when the current is switched off. Thus the single prolonged pulse will be converted into two separate pulses of opposite polarity.

Another indication that this might be the case is that the very accurate work of Bourguignon and Déjean (1926) and of Achelis and Merkulow (1930) using condenser shocks, did not show the raised threshold for longer pulses. This may be because condensers produce a pulse which does not switch off sharply but which dies away exponentially. There will be no "off" effect of the kind which might be found with constant current or constant voltage stimulation with square pulses. This made it

seem likely that the sharp "off" effect was responsible for the complexity of the curve.

Tests of the Hypothesis

(i) The next experiment tested this hypothesis by reproducing the postulated conditions. A 7-millisecond anodal pulse, 80 per cent. of the threshold intensity, was followed after a varied interval by a 10-millisecond cathodal pulse whose intensity could be continuously varied. The hypothesis predicts a minimal threshold when the two are 30 milliseconds apart. This prediction is completely confirmed. In fact the whole curve under these conditions follows closely the strength-duration curve for a single pulse. (Table II, Figure 1 B).

So far no variation in the curve has been detected when the polarities of the two pulses are exchanged, nor when the order of the fixed and variable pulses is changed. The threshold at 30 milliseconds is always less than the threshold for the 10-millisecond pulse alone, and at 70 milliseconds is always greater.

(ii) Although this explanation may hold good for periodic stimuli, it may still not be wholly true for the response to continuous stimulation. To see whether there really is an "off" response after continuous stimulation another experiment was devised. A subthreshold pulse of 75 milliseconds duration was followed at a varied interval by a 10-millisecond pulse of 50 per cent. greater intensity. The present hypothesis predicts three separate "on," "off," "on," processes; the interval between the last two alone is varied and one must predict a minimum threshold when they are 30 milliseconds apart. In fact a minimum is found at about 40 milliseconds. (Table III, Figure 1 C).

This discrepancy is not great and may be attributed to the disturbing effect of the first immobile "on" response. Further, all other hypotheses fail to predict any minimum at all. Any slow inhibitory process ought to be well under way after 75 milliseconds. Alternatively, if this were simply the same kind of situation as that shown in Figure 2 a steady rise in the threshold would be expected. This seems then fairly conclusive evidence that separate "on" and "off" processes are aroused, that they interact, and that they are of opposite polarity. If this is the case, they are very similar to the "on" and "off" processes already known as the retina's response to light.(Kuffler, 1953; Gernandt & Granit, 1947; Barlow, 1953; Thomson, 1953; Hartline, 1938; Granit, 1947), and they are in fact almost certainly the same (see Discussion). Granit (1946) has produced experimental evidence of the opposite electrical polarity of the "on" and "off" elements.

All the previous work on the "on" and "off" fibres has been done with microelectrodes and their interaction has not been fully investigated because of the technical difficulty of the experiments (see Thomson, 1953). Their temporal characteristics are quite unknown, nor do we know in what sense they interact. The experiments seem to show that, in the retina at least, they are always antagonistic. (Kuffler, 1953). But this applies only to prolonged and simultaneous stimulation of adjacent areas of the retina. The present series of experiments seems to show that the interaction may be complex and important, and, by showing the time course of the interaction, provides a good grounding for further experimental investigation of the problem.

Site of Stimulation

The present gross experiments do not allow the separation of the different possible levels of interaction. The "on" and "off" fibres could interact beyond the optic nerve either in the thalamus or in the cortex. Alternatively, the interaction

could be taking place within the retina itself at a preganglionic level. The site of stimulation is very important in this connection. If the point of stimulation is the optic nerve (as maintained by Baumgardt) or the optic centres of the brain (as maintained by Copolongo) the interaction could not possibly take place within the retina. However the following experiments seem to eliminate both these possibilities. (The first experiment was suggested by Dr. F. W. Campbell of the Oxford Eye Hospital.)

(i) Pressure on the eyeball by occlusion of the retinal circulation leads to a darkening of the field of vision. This has been shown to be due to a temporary interruption of nervous conduction within the retina, leaving the photosensitive layer and the optic nerve itself intact. (Craik, 1940). The electrical threshold under these conditions is very much raised and does not seem to have the characteristic threshold-time curve of the normal condition (2 subjects). This is a most unpleasant experiment to do, so naturally not many readings were taken for each subject. However the threshold was raised by a factor of ten, so that the difference found is certainly reliable. The best explanation of this result is that under normal conditions it is the retina itself which is stimulated but when conduction through this is blocked, the optic nerve fibres react, but at a higher threshold.

(ii) In subjects with detachment of the retina a very similar disability is found which again is largely confined to the retina. They do have also some degeneration of the optic nerve but this is not usually very great. Two subjects were used and in both cases results almost identical with the above were found. The threshold was raised by a factor of ten and no sign could be found of any minimum in the strength duration curve. The chronaxie was probably of the order of a few milliseconds similar to that found for an ordinary motor or sensory nerve (see page 56). Here again the best explanation seems to be that it is the optic nerve which is being stimulated.

(iii) Despite the results of Copolongo (1948), blind subjects with atrophy of the optic nerve do not apparently experience the phosphene at any bearable intensity of stimulation (2 subjects). One of the subjects was used to electrical shocks in his work as a physiotherapist and was prepared to stand a current of 8 milliamps. He was still unable to see the phosphene although this is a frighteningly high current. Thus stimulation must normally be taking place below the level of the thalamus. The only likely explanation for Copolongo's results is that he was giving his patients such a severe shock that they saw "stars."

(iv) Another suggestive result is that in the dark adapted eye the strength-duration curve does not rise for long durations (5 subjects, Figure 3). It would be rash to offer a complete explanation of this, but it is obvious that the biggest changes in dark adaptation occur in the retina itself, so that here again a retinal change has resulted in a big change in the excitability. The mechanism of this change is probably related to the finding of Motokawa and Iwama (1949) and of Bouman (1935), that the ability of the eye to adapt to an electrical current is very much decreased in the dark adapted eye. If this is the case, there would be no "off" response when the current is switched off. As a result the threshold for short pulses is raised because of the absence of summation, and the threshold for long pulses is lowered because of the absence of any inhibitory "off" process. This is also in line with the established connection between the "on" fibres and the rods.

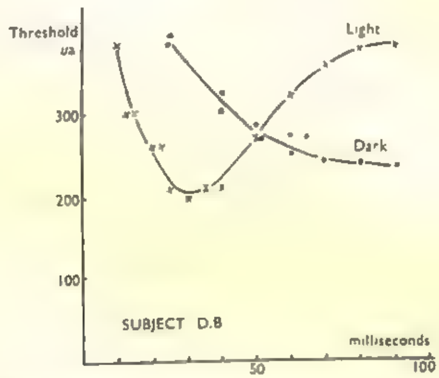
Incidentally these results may go a long way towards settling the controversy about the effect of dark adaptation on the electrical threshold. There is first of all a very great deal of individual variation. But a common result is the one shown

in Figure 3 where the threshold is raised in the dark for short pulses but lowered for longer ones.

These four experiments make it almost certain that the site of stimulation is the retina. The work of Bogoslowski and Ségal (1947, a & b) has already supported this. They showed that the spatial characteristics of the phosphene are what would be expected from the spread of current through the retina along the path of least resistance. These spatial characteristics of the phosphene would be hard to understand if the site of stimulation were anywhere else.

There is nothing in any of the above experiments to show where the interaction between the "on" and "off" processes takes place. They do show however that since the site of stimulation is in the retina, the interaction could also be taking place there, and the time relations found could be characteristic of the retina itself.

FIGURE 3.



Threshold as a function of time for single pulses in the light adapted eye and in the completely dark adapted eye. Similar results have been obtained from 4 other subjects.

The Accuracy of the Method

Because of previous disagreement among workers in this field, it would be undesirable to add yet another new result without some discussion of the research technique.

The results of this investigation are extremely repeatable. It was not felt necessary to use any function like Motokawa's ζ to eliminate the effect of changes in skin resistance. This was thought to be due to the use of constant current stimulation. Many of the previously reported extremely varied results may have been due to the use of voltage measures of threshold or of rather inefficient current generators.

Even if the results are extremely repeatable, they may still be inaccurate or misleading, because of distortion of the electrical pulse as it passes through the skin and other tissues surrounding the eye. There is evidence that such distortion does occur over a period of about 10 milliseconds from the beginning of the pulse. The big disagreements, in the previously reported values of the chronaxie, are due to this effect. Workers using condenser shock stimulation have obtained chronaxies of 1 to 2 milliseconds (Achelis and Merkulow, 1930); Bourguignon and Déjean, 1926). Others using constant current stimulation have found chronaxies of about 10 milliseconds. (Verrijp, 1924; Iwama, 1949). As Baumgardt has pointed out (1951a), the chronaxie concept is not applicable to strength duration curves with a trough

in them. He has used a similar measure, the "pseudochronaxie," to allow the same kind of comparison to be made with the results of other workers. Both he and the present author find "pseudochronaxies" of the same order as the chronaxies of Verrijp and Iwama.

The short chronaxies found using condenser shocks are thought to be due to the electrical capacity of the skin in series with the electrodes. If a constant voltage pulse is passed through a condenser the initial current will be very great, but will fall off rapidly as the opposing potential is built up. As a result, the longer pulses will be relatively ineffective, the rheobase will be raised and the chronaxie consequently shortened. If a condenser shock is used the effect will be still greater and the chronaxie still shorter. This argument leads to the specific prediction that the chronaxie found using constant voltage will be intermediate between the extremes with constant current and condenser shocks. This prediction was confirmed on two subjects for whom "pseudochronaxies" of about 6 milliseconds were obtained.

The effect amounts to a distortion of the constant voltage pulse by the series condensers of the skin. Constant current pulses can not be distorted by series condensers but could be distorted by parallel condensers. The effect could be either a raising or a lowering of the relative value of the rheobase with consequent shortening or lengthening of the chronaxie. Chronaxie is therefore a meaningless term in connection with the electrical stimulation of the eye through the skin.

Distortion is not appreciable over the range (10 to 100 milliseconds), of the present experiments. This is demonstrated in three different ways.

(1) The shape of the curve is essentially the same under constant voltage conditions.

(2) The shape of the curve is not changed when contact lens electrodes are used. Only one-tenth the amount of current is needed under these conditions, showing that at least nine-tenths of the current is by-passing the retina under the ordinary conditions of stimulation.

(3) When the voltage developed across the electrodes is applied to the plates of a cathode ray oscilloscope it is found to reach a maximum within 10 milliseconds. This shows that all the tissue condensers are fully charged within that time. Beyond this the tissues will act as a resistance network and will not distort the pulse.

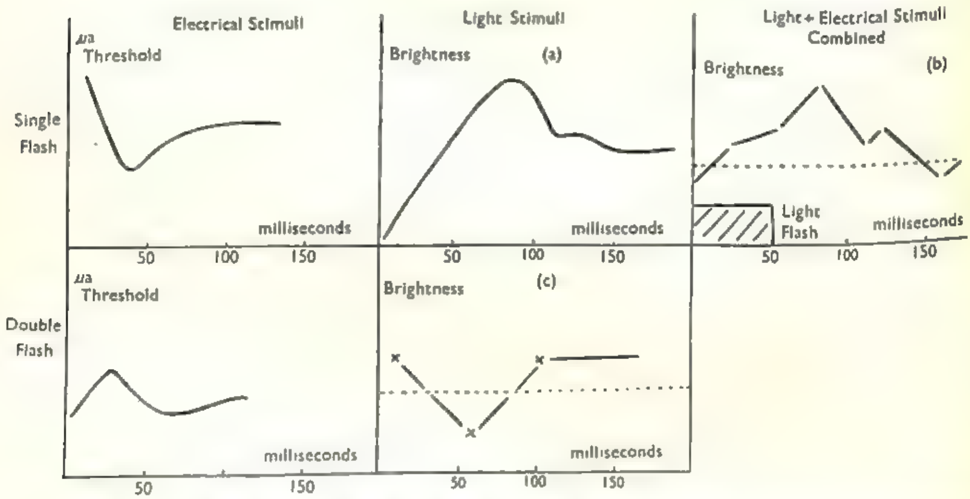
IV

Discussion

If the above arguments are valid, and there really is an interaction between "on" and "off" processes, one would expect the same mechanism to operate in response to light stimulation. There is no evidence for such a mechanism to be found in any of the very accurately carried out light-threshold experiments. This is probably because the slowness of the photochemical processes prevents the development of sharp "on" and "off" responses to threshold light stimulation. However when the intensity of the light is increased the photochemical reactions are faster, and sharp "on" and "off" processes can occur (Granit, 1947, p. 151). Thus although no light threshold experiments show these time relations a whole series of brightness comparison experiments do. The minimum in the threshold curve at 30 milliseconds corresponds to the well known (but rather neglected) Broca and Sultzer (1902) phenomenon (Figure 4). In these experiments it was shown that at a given intensity a light flash of only 60 milliseconds appears brighter than one lasting 200 milliseconds. Baumgardt and Segal (1946b) have shown that this difference affects the

reaction time and claim that this shows the "retroactive" nature of their slow inhibitory process. Comparable results have been found by Bogoslowski and Segal (1945, Figure 4) for the interaction of light and electrical current. In this experiment the effect of an electrical current through the eye on the brightness of a 50-millisecond flash of light was assessed on a 7-point subjective scale. No comparison light was used. As a result this experiment is not so accurate as that of Broca and Sulzer, but it is accurate enough to show that very much the same kind of curve is obtained and also to show that the same electrical pulse may be excitatory or inhibitory depending on the time between it and the light flash.

FIGURE 4.



Comparison of strength-duration curves for various forms of light and electrical stimulation.

(*a*) *Broca & Sulzer.* 1902. The brightness of a single flash of light of duration "t."
(*b*) *Bogoslowski & Segal.* 1945. The change in the brightness of a 50 millisecond flash produced by the "make" of a cathodal current through the eye at time "t."
(*c*) *Baumgardt & Segal.* 1946 (a). The fused brightness of two 10 millisecond flashes of light separated by time "t."

Baumgardt and Segal (1946, see Figure 4) performed a double-flash experiment with light. They used two 10-millisecond superimposed flashes of light. The first flash was smaller and within the area of the second. When they were only 10 milliseconds apart the central area appeared lighter than the rest. When they were 50 milliseconds apart it appeared dark. At about 100 milliseconds the two flashes appeared separate and the central area was bright again. This result is comparable with Baumgardt's (1951b) experiment on the threshold for two 10-millisecond electrical pulses.

All these experiments have been discussed by Baumgardt and Segal (1947) and their theory applied to them. The present author's hypothesis is equally capable of explaining them.

The evidence is not yet sufficient for much to be said about the site of the interaction, or the reasons for the peculiar time relations which are found. There are however two attractive pieces of speculation in the literature which should perhaps be discussed. The first is to be found in Adrian's "Basis of Sensation" (1928).

Here the Broca and Sulzer curve is presented side by side with the Adrian and Matthews curve, showing the frequency of impulses in the optic nerve of the conger eel following the onset of a flash of light. From the similarity in the two curves one is presumably expected to infer a similarity of mechanism. If this comparison were valid the site of interaction would be placed firmly in the retina. But since one curve shows the intensity of the response as a function of the time relations of the stimulus while the other gives it as a function of the time relations of the response, these curves cannot be directly compared. There is certainly no reason to expect them to be due to the same mechanism.

The second is the suggestion by Bartley (1939) that the enhanced brightness of a slowly flashing light (Brücke effect) is due to the driving of the alpha rhythm by the stimulus. If the response to a single flash of light is double, as is suggested in this paper, it is conceivable that such a driving effect could take place for the single flash as well as for repeated flashes. This would be a particularly attractive theory because it would account for the effects of changing the polarity of the second of the two pulses in the double pulse experiments. Unfortunately although the times found in the experiments with light are of the right order, the 30-millisecond minimum is outside the known range of the alpha rhythm. The same kind of theory might still be used if a centre with the appropriate periodicity could be found in the visual system. There are other indications that this sort of "resonance" theory would be useful. For example, Motokawa and Iwama (1950) have found several "resonant" minima in the strength-frequency curve for A.C. stimulation of the eye and Bartley (1951) has shown that the apparent frequency of a just perceptibly flickering light is always about 18 cps. The present author has evidence that the frequency of even quite strongly flickering phosphenes always appears to be about 18 cps. So far there is no reason to think that the retina has a spontaneous rhythm of this frequency, so if this kind of explanation is the right one some central system must be supposed to be the site of the interaction. There are however other possible mechanisms and a full discussion of them should be left until more evidence is available.

The longer time constants found in the experiments with light are puzzling. They seem to be due in some way to the "blurring" of the stimulus by the photochemical process. This is shown by the shortening of the time constants as the intensity of the light is increased. With increasing intensity of the light stimulus the electrical events at the retina will become sharper in onset and in form. The more rapid these events become the more they will approach the ideal form which is the result of square wave electrical stimulation.

It is worth noting that one of the simplest ways of explaining the effects discussed in this paper would be to postulate interaction in the retina of the electrical potentials which develop in response to both light and electrical stimulation. If the "off" response were an inverted form of the triphasic "on" response recently found by Cobb and Morton (1952), then the interaction of the "on" and "off" responses would be of exactly the type required. The interaction of two triphasic waves inverted in the m–w sense would give the required succession of inhibition–summation–inhibition as the amount of overlap between them is decreased. This amounts to a prediction of the form of the "off" response in the human retinogram which the author intends to investigate.

ACKNOWLEDGEMENTS

The above experiments were carried out while receiving a grant from the Medical Research Council.

I wish to thank Professor G. Humphrey for providing laboratory facilities, Dr. R. S. Creed for many valuable discussions of the problems encountered and for reading the manuscript, and Mr. J. P. F. Lloyd for providing blind subjects. I am particularly indebted to my supervisor, Mr. J. A. Deutsch, for help and encouragement at all stages of the investigation.

The electronic stimulator was constructed by Mr. W. J. Markwick.

References

ABE, Z. (1950). Influence of adaptation on the strength-frequency curves of human eyes, as determined by the electrically produced flickering phosphene. *Tohoku J. exp. Med.*, **54**, 37–44.

ACHELIS, J. D., and MERKULOW, J. (1930). Die electrische Erregbarkeit des menschlichen Auges während der Dunkeladaptation. *Z. Sinnesphysiol.*, **60**, 95–125.

BARLOW, H. B. (1953). Action potentials from the frogs retina. *J. Physiol.*, **119**, 58–68.

BARLOW, H. B., KOHN, H. I., and WALSH, E. G. (1947). The effect of dark adaptation and of light on the electrical threshold of the human eye. *Amer. J. Physiol.*, **148**, 376–381.

BARNETT, A., and POSNER, M. (1941). Effects produced by impressing alternating electrical currents on the human skull. *J. Psychol.*, **11**, 271–278.

BARTLEY, S. H. (1939). Some factors in brightness discrimination. *Psychol. Rev.*, **46**, 337–358.

BARTLEY, S. H. (1951). In: *Handbook of Experimental Psychology.* Ed. Stevens. New York.

BAUMGARDT, E. (1951a). Sur le seuil du phosphène électrique Quantité liminaire et pseudochronaxie. *C.R. Soc. Biol., Paris*, **145**, 1654–1657.

BAUMGARDT, E. (1951b). Sur le seuil du phosphène électrique. L'inhibition rétroactive révélée par la méthode du double choc. *C.R. Soc. Biol., Paris*, **145**, 1824–1826.

BAUMGARDT, E., and SÉGAL, J. (1946a). Le fonction inhibitrice dans le processus visuel. *C.R. Soc. Biol., Paris*, **140**, 231–233.

BAUMGARDT, E., and SÉGAL, J. (1946b). L'inhibition rétroactive dans le vision demontrée par la méthode des temps de réaction. *C.R. Soc. Biol., Paris*, **140**, 267–269.

BAUMGARDT, E., and SÉGAL, J. (1947). Facilitation et inhibition, paramètres de la fonction visuelle. *Année psychol.*, **43–44**, 54–102.

BOGOSLOWSKI, A. I., and SÉGAL, J. (1945). Facilitation et inhibition dans le phosphène électrique. *C.R. Soc. Biol., Paris*, **139**, 290–291.

BOGOSLOWSKI, A. I., and SÉGAL, J. (1947a). La sensibilité électrique de l'oeil. 1re partie. La phenomenologie du phosphène électrique. *J. Physiol. Path. gén.*, **39**, 87–99.

BOGOSLOWSKI, A. I. and SEGAL, J. (1947b). La sensibilité électrique de l'oeil. 2me partie. Analyse des facteurs physique et physiologique dans l'excitabilité électrique de l'organe visuelle. *J. Physiol. Path. gén.*, **39**, 102–117.

BOUMAN, H. D. (1935). The electrical excitability of the eye. *Arch. néerl. Physiol.*, **20**, 430–445.

BOURGUIGNON, C., and DÉJEAN, R. (1926). Double chronaxie du système optique de l'homme. Lois d' Hoorweg., *C.R. Soc. Biol., Paris*, **94**, 750–753.

BROCA, A., and SULZER, D. (1902). La sensation lumineuse en fonction du temps. *J. Physiol. Path. gén.*, **4**, 632–640.

CRAIK, K. J. W. (1940). Origin of visual after images. *Nature*, **145**, 512.

COBB, W., and MORTON, H. B. (1952). The human retinogram in response to high intensity flashes. *E.E.G. clin. Neurophysiol.*, **4**, 547–556.

COPOLONGO. (1948). *Rivista oto-neuro-oftalm.*, **23**, 371.

GEBHARD, J. W. (1953). A revue of Motokawa's work on the electrical stimulation of the human eye. *Psychol. Bull.* **50**, 73–111.

GRANIT, R. (1946). Distribution of Excitation and Inhibition in Single Fibre Responses from polarised Retina. *J. Physiol.*, 1946, **105**, 45–53.

GRANIT, R. (1947). *Sensory Mechanisms of the Retina.* London.

GERNANDT, B., and GRANIT, R. (1947). Single fibre analysis of inhibition and polarity of retinal elements. *J. Neurophysiol.*, **10**, 295–301.

HARTLINE, H. K. (1938). The response of single nerve fibres of the vertebrate eye to illumination of the retina. *Amer. J. Physiol.*, **121**, 400–415.

IWAMA, K. (1949). Die electrische Erregbarkeit des menschlichen Auges. *Tohoku. J. exp. Med.*, **50**, 71–77.

KUFFLER, S. (1953). Discharge patterns and functional organisation of mammalian retina. *J. Neurophysiol.*, **16**, 37–68.
MOTOKAWA, K. (1949a). Retinal processes and their role in colour vision. *J. Neurophysiol.*, **12**, 291–303.
MOTOKAWA, K. (1949b). Physiological studies on mechanisms of colour reception in normal and colour blind subjects. *J. Neurophysiol.*, **12**, 465–474.
MOTOKAWA, K. (1949c). Physiological induction in human retina as the basis of colour and brightness contrast. *J. Neurophysiol.*, **12**, 475–488.
MOTOKAWA, K. (1950). Field of retinal induction and optical illusions. *J. Neurophysiol.*, **12**, 413–426.
MOTOKAWA, K. (1951). Propagation of retinal induction. *J. Neurophysiol.*, **14**, 339–351.
MOTOKAWA, K., and EBE, M. (1952). Selective stimulation of colour receptors with alternating currents. *Science*, **116**, 92–94.
MOTOKAWA, K., and IWAMA, K. (1949). Electrical stimulation of the human retina with exponentionally increasing currents. *Tohoku. J. exp. Med.*, **50**, 25–37.
MOTOKAWA, K., and IWAMA, K. (1950). Resonance in the electrical stimulation of the eye. *Tohoku J. exp. Med.*, **53**, 201–206.
MOTOKAWA, K., IWAMA, K., and EBE, M. (1952). Retinal colour processes in cats. *Jap. J. Physiol.*, **2**, 198–207.
MOTOKAWA, K., IWAMA, K., and TUKAHARA, S. (1951). Colour processes in single retinal elements. *Tohoku J. exp. Med.*, **53**, 399–406.
POLLOCK and MEYER. (1938). An adaptation-like phenomenon of electrically produced phosphenes. *Amer. J. Physiol.*, **122**, 57.
ROHRACHER, H. (1935). Über subjectiver Lichterscheinungen bei Reizung mit Wechselstromen. *Z. Sinnesphysiol.*, **66**, 164–181.
SCHWARZ, F. (1947). Über die electrische Reizbarkeit des Auges bei Hellund Dunkeladaption. *Pflüg. Arch. ges. Physiol.*, **249**, 76–86.
THOMSON, L. C. (1953). The localisation of function in the rabbit retina. *J. Physiol.*, **119**, 191–209.
VERRIJP, C. D. (1924). L'influence de l'adaptation a l'obscurité sur l'excitabilité électrique de l'oeil humain. *C.R. Soc. Biol., Paris*, **92**, 55–58.

RESPONSE—DURATION OF LEVER PRESSING IN THE RAT

BY

H. M. B. HURWITZ

(*From the Animal Laboratory, Birkbeck College, London*)

Two aspects of the lever-holding behaviour of rats in a Skinner-box have been analysed: firstly, the changes in the duration of the responses during the acquisition and experimental extinction phase, and, secondly, the bunching of responses during the experimental extinction phase. The response-durations on the first few acquisition trials were found to be bimodally distributed, but to become stabilized at 0·42 seconds as practice increased. During experimental extinction response-durations increased. The rate of increase depended on the conditions of secondary reward. Rats who had the source of secondary reward removed, or restricted, yielded a steeper slope than those who responded under the same conditions of secondary reward as during training.

The analysis of the extinction scores revealed that unrewarded responses were emitted in groups. Response-group latencies were shown to progress as a positive function of the number of response-groups, and the average response-duration for consecutive response-groups to increase progressively. The slope of the duration curve plotted for successive response-groups increased, and this was related to an increase in the value of the time-intercept for successive response-groups.

I

Introduction

Hungry laboratory rats are placed in a Skinner-box in which food is secured by pressing a light, protruding lever. When activated the food magazine emits a sharp, audible click. It has been found that during a test period in which no food is delivered, animals hearing the click will make more responses than those who have this source of stimulation removed. The click is said to have become a secondary reward.

The present paper describes an experiment which showed that the duration of the lever-holding response increases during an extended period of unrewarded responding. The rate of the change in the duration was found to be related to the prevailing secondary-reward conditions. Attention is also drawn to the emergence of response-groups, and an attempt is made to relate response-groups to response-durations. These findings shed further light on the character of the rat's behaviour in the Skinner-box and the factors which determine it.

Skinner and others had previously noted that unrewarded responses occur spasmodically, at irregular intervals, and have maintained that this indicates agitated behaviour produced by an emotional effect which is set up through "lack of reinforcement" of a rigorously built-up behaviour chain. This finding runs counter to the view that responses are emitted at a constant rate, and confounds attempts to describe the rate of lever-pressing in terms of smooth, negatively accelerated curves. It obliges us to analyse the phenomenon of response-grouping more closely. The present paper includes some suggestions as to how this objective may be accomplished.

II

Experimental Design, Apparatus and Procedure

The experimental design of this study is detailed in Table I. There were five experimental groups of three months old, hooded male rats. The experimental conditions during learning and experimental extinction trials are outlined in Table I. *Group A* (7 animals) received a buzz which lasted for the total duration of each rewarded response

but had this source of stimulation removed during unrewarded responding trials. *Groups B and C* (6 and 7 animals respectively) received a buzz for approximately 0·5 seconds irrespective of the duration of the rewarded response, though Group C received a buzz for the total time the lever was depressed during unrewarded responding as against 0·5 seconds for Group B. *Group D* (4 animals) had the procedure for Group C reversed, i.e. a buzz for duration of rewarded response and 0·5 seconds for each unrewarded response was given, whilst *Group E* (4 animals) obtained a buzz for the total period of lever-pressing during rewarded and unrewarded responding.

TABLE I

Group	*No. of animals*	*Number of rewarded responses*	*Phase I Stimulus conditions during rewarded responses*	*Phase II Stimulus conditions during unrewarded responses*
A	7	90	0·5 seconds buzz	no buzz
B	6	90	0·5 seconds buzz	0·5 seconds buzz per response
C	7	90	0·5 seconds buzz	buzz for duration of response
D	4	90	buzz for duration of response	0·5 seconds buzz per response
E	4	90	buzz for duration of response	buzz for duration of response

Apparatus

The apparatus used in this experiment was a modified Skinner-box, 1 ft. in all dimensions. The modifications were as follows: (*a*) the lever, a 1 in. square brass plate protruding 1¾ in. into the box, was fixed 5 in. to the right of a semi-circular food trough. This forced the animal to shuttle between lever and trough. A pressure of 5 gms. was required to depress the lever. (*b*) A new pellet-delivery machine, described in detail in a previous paper (Hurwitz, 1953) was used in this study. Though the machine is almost noiseless the buzz has been introduced to cloak any noise emitted by it. (*c*) The apparatus was permanently fixed in a sound-resistant case.

Records

Lever-pressing responses were recorded on a paper tape moving at 1 cm. per sec. This device presented the data in terms of (*a*) the *rate* of responding, (*b*) the *duration* of each response and (*c*) the *time interval* between consecutive responses.

Procedure

The experimental procedure followed the customary pattern. For the first five days the animals were fed once every 22 hours on wet mash and cod-liver oil. On the 5th day of the feeding schedule the rat was allowed 20 minutes exploration in the apparatus after the customary 2-hours feeding. The following day it was placed in the apparatus before normal feeding, and allowed to eat 10 pellets of food. Each pellet weighed approximately 0·05 gms. and measured ¼ in. × ⅛ in. The animal was then withdrawn from the apparatus and returned to its feeding cage. On the following day it was once more introduced into the apparatus before the 2-hour feeding session. Ten pellets lay in the food trough. After the 7th pellet had been eaten, the experimenter delivered 30 pellets into the trough separately by operating an external switch which activated the feeding machine. Each pellet was accompanied by a 0·5 second buzz.

On the seventh day, the lever was connected to the food delivery machine. The animal was retained in the apparatus until it had obtained 30 pellets by pressing the lever. The training session was followed by an hour-and-a-half feeding in the cage.

A further learning session involving 60 pellets followed on the next day. On the ninth, and final day, no pellets were inserted into the food-magazine of the feeder. Unrewarded responding was continued until a period of 4 minutes elapsed during which no lever responses were recorded.

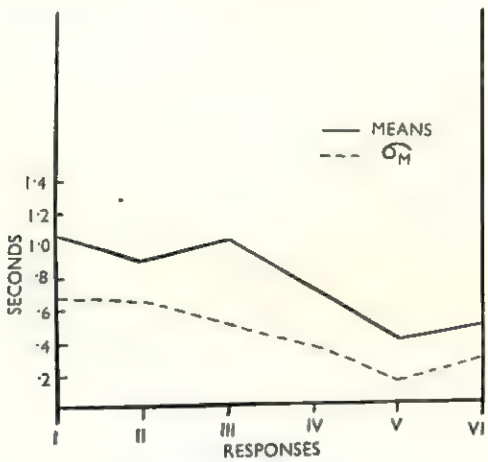
III

Analysis of the Results

Response-duration during rewarded responding

(*a*) In order to determine whether the duration of the response underwent a change as learning proceeded, the first 30 rewarded lever-responses made by each of the rats were analysed. To facilitate statistical treatment, the responses were grouped into units of 5, and the means computed. Figure I summarizes the results.

FIGURE I.



(in units of 5)

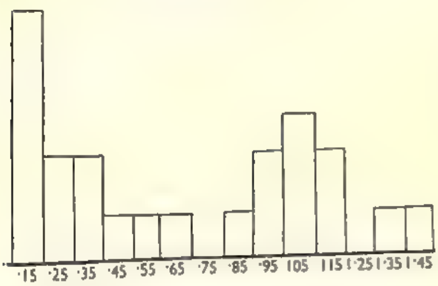
The unbroken line represents the means of the response-durations for all 28 animals. The broken line represents the standard deviation from the means. Both curves are decelerated, with the means stabilizing at 0·42 seconds. As a check on the reliability of this lower limit, a similar analysis was made for responses 30–60 and 60–90. This showed that the mean had increased to 0·55 seconds though the former analysis is in close agreement with the means computed for the first 30 responses, 0·47 seconds. The behavioural fact here described, that the duration scores of massed responses is an accelerated function of the number of responses made may well be due to a complicity of factors such as fatigue, satiation and the concomittant diminution of drive-intensity, and the general conditions of responding. This phenomenon has been more thoroughly studied during the past few months and the results of this investigation will become available in due course. Figure I also illustrates the decrease in the variability of the mean scores (broken lines). Further reference to this aspect of the data is made in section (c).

(*b*) It will be remembered that the five experimental groups fall into two groups with respect to the secondary-reward conditions obtaining during the first phase of the experiment. Groups A, B and C received a buzz lasting approximately 0·5 seconds for each lever response made, as against Groups D and E for which the buzzer sounded for as long as the lever was held down. The analysis of this data

suggests that the difference in the secondary-reward conditions during learning and rewarded performance is not reflected in the duration scores. Thus, on this showing, duration-score is not a sensitive measure for detecting the influence of these experimental conditions.

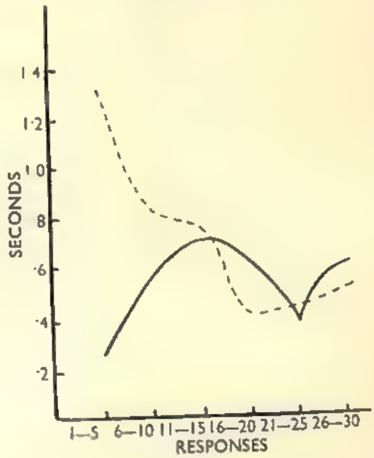
(*c*) The most significant changes in the standard deviation scores occur within the first 20 responses. This is also borne out by the coefficient of variation which

FIGURE 2.



Histogram of Response Duration Scores 1–15 (Means +)

FIGURE 3.



shifts from 75 per cent. to 60 per cent. This may be accounted for in the following manner: let us assume that the first lever responses were spaced, and separated by other activity, including trough exploration. Finding food has three immediately apparent effects: (i) the animal acts more vigorously; (ii) if the animal's food seeking activity has previously flagged, or given place to, for instance, escape behaviour, the food seeking behaviour regains its dominance; (iii) the likelihood of a repetiton of those actions (possibly accidental lever pressing) which preceded the finding of food, is considerably increased. The consistency of the change in the duration-scores of the first set of 30 responses may therefore be attributed to a period of "experimenting," leading gradually to the learning of relevant cues and appropriate actions. Some writers, notably Tolman and Krechevsky, would describe this behaviour in terms of changes in "hypotheses" as to "what leads to what." But this view does not take account of two distinguishable stages which may be described as, firstly, a period in which the individual learns that it has to perform a "ritual" before food is found in the trough; and, secondly, a period in which the "ritual" is perfected (technique or skill is developed). Only the former bears favourable comparison with the "formation of hypotheses," the setting up of "expectations" or "superstitions."

(*d*) In order to gain a more comprehensive view of the changes in the variability of the scores, a further analysis was carried out. The mean of the first 15 responses made by each animal was computed, and a frequency distribution drawn. The distribution has been reproduced in Figure 2. It strongly suggests a bimodal distribution, the modes falling between 0·20 to 0·29 seconds and 1·10 to 1·19 seconds respectively. If this histogram is considered in conjunction with the data displayed in Figure 1, it appears that there are two distinct populations: the initial duration

progressing from about 0·25 seconds in one case and 1·3 seconds in the other, to about 0·42 seconds between the 18th and 27th trials in both cases (Figure 3). This bimodality may be accounted for by the fact that the majority of animals encountered the lever in one of two ways: they were using the lever as a support for scaling the walls, or they hit the lever with the front paws whilst investigating the corner in which it had been mounted. In the former case, response-duration would be large, in the latter, small. However, the graph shows that irrespective of the initial "hypotheses" formed, the rats ultimately adopted similar modes of approach.

Response duration during unrewarded responding

Experimental extinction scores, i.e. measures obtained during a period when the customary reward has been withdrawn, have been extensively used to determine the degree to which learning had been accomplished under the conditions set up by the experimenter. Many writers have assumed that both the rate of responding (r) and the total number of responses made within an arbitrarily selected period (n) are the most representative forms of extinction scores, in that they act as reliable measures of the effects of such major determinants of behaviour as motivation and previous learning (Hull, 1943, 1951; Perin, 1942; Skinner, 1938). Skinner, for instance, has recently repeated his claim that "the rate of responding appears to be the only datum that varies significantly and in the expected direction under conditions which are relevant to the 'learning process'." (Skinner, 1948). But the results of our experiments suggest that these claims are excessive: both the rate of responding and the total number of responses made within a selected period of observation are very rough guides to the changes in behaviour occurring during unrewarded responding and to the changes which had occurred throughout the learning period.

(*a*) *Rate of responding:* Rate of responding may be directly and indirectly recorded. The former method requires a cumulative kymograph record to be made, from which the slope of the curve may be immediately computed (Skinner, 1938). The indirect method, on the other hand, has been favoured in experiments employing large groups of animals (Ellson, 1938; Williams, 1938). The time between the first and final response made within the session is divided into suitable units and the sum of the number of responses occurring within each period found.

The assumption underlying the indirect method is that the trend exhibited in the group curve is a "*generalized*" representation of the empirical data. It is sometimes assumed that the rate of responding is a continuous function of time (t) as, for example, in calculating reaction latencies, "str." This assumption obscures much interesting information in the data. As already noted, even in the earliest reports of the rat's lever pressing behaviour, it was remarked that unrewarded responding is spasmodic, rather than continuous, in character; that is, unrewarded responses occur in "bursts." However, except for occasional passing references, no attempt to analyse this phenomenon has so far been made. It is conceivable that these bursts of responses, or *response-groups*, follow a lawful pattern. Should this be the case, the analysis would be of considerable importance for the construction of curves, or the finding of mathematical functions, purporting to give a realistic description of the animal's behaviour. For instance, it would suggest the construction of two curves, one representing changes in the *rate of responding within* the *response-groups* (r_{RG}), the other describing the *rate at which response-groups emerge* (r_o).

(*b*) *Total number of unrewarded responses:* The total number of unrewarded responses (n) made by the animal within a set period represents the alternative method of measuring. In recent years it has become common practice to employ

a criterion of 4 to 5 minutes of no-response for terminating this phase of the experiment (Ellson, 1938; Koch, 1945; Perin, 1942; Saltzman and Koch, 1948; Williams, 1938). It follows from the considerations put forward earlier in this paper that this criterion fails to distinguish between the rate at which response-groups emerge as well as the rate of emittance of responses within each response-group. These two measures taken together may be more sensitive than the overall measure of "number of responses made to an extinction-criterion."

Moreover, it is notoriously difficult to obtain statistically significant differences between experimental groups in a Skinner-box when using "n." Even with large groups of animals, as in the Yale studies, the dispersion scores remain large. Furthermore, curves drawn from median and mean total scores frequently differ markedly from each other; different mathematical functions are needed for each and this is quite unpredictable, theoretically. The gross variability in the data may rest largely in the experimental techniques employed, and not in the experimental subjects.

It is possible to reduce the inconclusiveness of this type of data by eliminating fortuitous factors and by elaborating techniques of measurement. The latter is attempted in the next section.

(*c*) *Response-duration of unrewarded responses:* Table 2 presents the results of the second phase of the experiment (unrewarded responding). The medians of the *number* of unrewarded responses, "n," made by each of the experimental groups

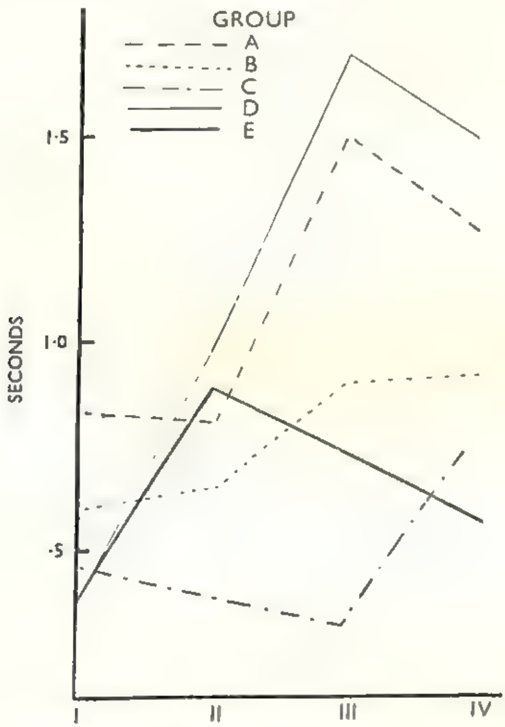
TABLE II

Group	*I* *Average response duration*	*II* *Unrewarded responding* — *No. of unrewarded responses (medians)*	*II* *Unrewarded responding* — *Secondary reward conditions*
A	0·92	19	Source of secondary reward removed
B	0·78	32	0·5 seconds buzz
C	0·50	25	buzz for duration of response
D	1·20	25	0·5 seconds buzz
E	0·66	29	buzz for duration of response

shows the usual degree of dispersion (column II). The difference between the original conditions of training and those prevailing during unrewarded responding were not detectable by this measure. However, the existence of a discernable difference is shown by the results detailed in column I. Rats who obtained a buzz lasting 0·5 seconds for each unrewarded response (Group B and D) and those (Group A) for whom the source of secondary reward had been removed, all obtained high duration-scores. In the former, the sigma was 0·05, yielding a coefficient of variation of 10 per cent.—a remarkable consistency. The duration-scores for each of the five experimental groups are shown in Figure 4. Curves A and B are somewhat similar, as are curves B, C and E. In order to interpret these results we draw attention to the fact that both Group A and B have undergone a form of stimulus

deprivation—the former more so than the latter. The remaining groups responded under secondary reward conditions which were either the same as during rewarded responding or which involved an increase of such stimulation. However, despite the similarity in the form of the curves A and D, the origin of the two functions, 0·85 seconds and 0·42 seconds, respectively—a difference significant at the 1 per cent. level—suggests the possible operation of an additional factor. A hint, making the results more intelligible, comes from considering the course of the animal's learning. The task set involves that the animal press the lever, shuttle to the food trough, retrieve the pellet of food and having eaten it return to press the lever. During the

FIGURE 4.



Unrewarded Responses (grouped in units of 5)

early trials, however, a number of "abortive" lever-pressing responses occur, i.e. the lever is not sufficiently depressed to activate the food delivery mechanism and buzzer. Nevertheless, the animal shuttles to the food trough, goes unrewarded and promptly returns to make a more vigorous, extended lever-response which is accompanied by the buzz and followed by the customary reward. Thus, the vigorous, extended lever-responding observed during the second phase of the experiment had previously been rewarded as against "lethargic," "buzzerless" efforts. In the case of Group A this learning immediately manifested itself during unrewarded responding sessions. The view herein implied, that the animal is pressing the lever in order to get food *and* the buzz is supported by a test made after the completion of the main experiment. Group A was readmitted to the Skinner-box. The buzzer was connected to the lever, as for Group C, but the food magazine remained inactive. The response durations during this phase of the experiment were remarkably like those for Group B, C and D, except that the rate was more accelerated.

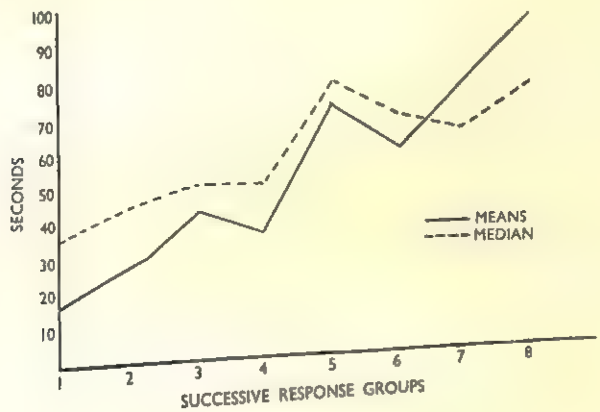
To conclude; the absence of the buzz accelerated the emergence of new behaviour patterns, such as more exaggerated lever-pressings. The results of Frick and Miller (1951) and unpublished results from our laboratory, on changes in the pattern of responding during rewarded and unrewarded trials, is relevant to the point raised here. It will be discussed in greater detail at a later date.

We furthermore predict that the duration-scores for the differing groups would attain the same upper limit if the unrewarded responding session would be extended to, say, an hour. The rate of approach to this upper limit, however, is dependent on distinctive secondary-reward conditions. It is apparent from this discussion of the changes in response-duration during both rewarded and unrewarded responding sessions, that more research must be done before the assumed relationships can be accommodated into systematic theory.

Response-groups

As already stated, unrewarded lever-responses occur in closely bunched groups. It must be left to subsequent research to investigate the specific properties of response-groups, and to determine the structural-functional properties of this phenomenon. The results here reported represent a first attempt to clarify some of the issues involved.

FIGURE 5.



(*a*) A criterion of 15 seconds between two consecutive lever-responses was selected in order to break down the data and isolate response-groups. The time intervals between response-groups, i.e. response-group latencies (t_o), were determined for all subjects and the mean and medians values found. The results are graphed in Figure 5. The intervals between the response-groups increase with the number of the response-group. These unrewarded responding sessions lasted until 4 minutes no lever-response. Subsequent studies have shown that in the type of Skinner-box employed in these experiments, a response recovery occurs shortly after the period laid down by this criterion.

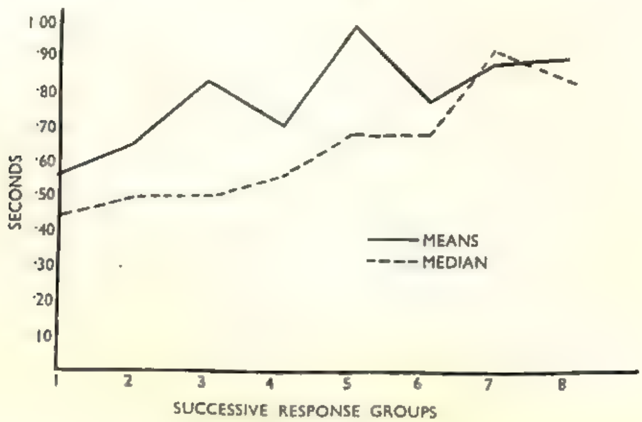
(*b*) It has previously been shown that the duration-scores increase as a positive function of the successive number of unrewarded responses made. In Figure 6, the mean and median duration-scores for response-groups have been plotted. It will be seen that the distribution of duration-scores become progressively less skewed,

i.e. the mean and medians are converging as a positive function of the order of response-groups. Three factors appear to determine this function:

(i) duration-scores within each response-group tend to increase;
(ii) the slope of the duration-curve tends to increase with successive groups;
(iii) the time-intercept value increases gradually with successive response-groups.

The specification of these three factors is tentative.

FIGURE 6.



(*c*) Finally, we draw attention to a number of problems concerning response-groupings which, in addition to those already considered in the preceding sections, require urgent investigation. These problems are:

(i) what factors affect the number of responses made in successive groups;
(ii) what factors affect the length of consecutive response-groups as measured by the interval between the first and final responses;
(iii) what factors affect the successive response-latencies in successive response-groups;
(iv) what is the relation between response-groups defined in terms of consecutive responses to the lever, or trough, as distinct from response patterns which involve making a series of heterogeneous responses.

IV

Summary

The present paper reports on an investigation of the factors determining changes in the lever-holding behaviour of rats in a Skinner-box. It was argued that the duration measure is sensitive to the following conditions:

(i) It is subject to systematic variation during the initial period of learning;
(ii) the variation decreases with increased practice (development of response-technique);
(iii) response-duration show a slight increase during extended rewarded practice, which may be related to the generation of fatigue, or a reduction in the motivation of the animal;
(iv) during unrewarded responding, response-duration increases. The rate of acceleration appears to be related to secondary reward conditions.

Stress has also been placed on the emergence during unrewarded responding of response-groups. Response-groups are defined as closely bunched, consecutive responses occurring within 15 seconds of one another. It was shown that respons-group latencies progress as a positive function of the number of response-groups. Furthermore, the following relations between response-groups and response-duration scores was indicated:

(i) the average response-duration for consecutive response-groups increases progressively;

(ii) the skewness in the response-duration distribution progressively decreases. This was due to an increase in the slope of the duration curve plotted for successive response-groups, with a corresponding increase in the value of the time-intercept for successive response-groups.

The author wishes to thank Mr. B. M. Foss and Mr. A. R. Jonckheere for invaluable advice in the preparation of this paper. Mr. John Whitfield kindly commented on the final draft and suggested several important changes in the presentation of the data.

The construction of the apparatus was made possible by a subsidy from the Central Research Fund, University of London. I am deeply indebted to Mr. L. Thomas for building the food delivery machine, and to Mr. R. Cohen for assistance in collecting the data.

Special thanks are due to Professor C. A. Mace and Mr. B. M. Foss for their constant encouragement, and for providing the facilities for this research.

References

Bugelski, B. R. (1938). Extinction with and without sub-goal reinforcement. *J. comp. Psychol.*, **26**, 121–134.

Denny, M. R. (1948). The effect of using differential end boxes in a simple T-maze learning situation. *J. exp. Psychol.*, **38**, 245–249.

Ellson, D. G. (1938). Quantitative studies in the interaction of simple habits. I. Recovery from specific and generalized effects of extinction. *J. exp. Psychol.*, **23**, 339–359.

Grice, R. G. (1948). The relation of secondary reinforcement to delayed reward in visual discrimination learning. *J. exp. Psychol.*, **38**, 1–16.

Hull, C. L. (1943). *Principles of Behaviour*. New York.

Hull, C. L. (1951). *Essentials of Behaviour*. Yale.

Hurwitz, H. M. B. (1953). A new rat-pellet feeding machine. *Quart. J. exp. Pssychol.*, **5**, 36.

Koch, S., and Daniels, W. J. (1945). The effect of satiation on the behaviour mediated by a habit of maximum strength. *J. exp. Psychol.*, **35**, 167–187.

Miller, N. E. and Stevenson, S. S. (1936). Agitated behaviour of rats during experimental extinction and a curve of spontaneous recovery. *J. comp. Psychol.*, **21**, 205–231.

Mote, F. A., Jnr., and Finger, F. W. (1942). Exploratory drive and secondary reinforcement in the acquisition and extinction of a simple running response. *J. exp. Psychol.*, **31**, 57–68.

Mowrer, O. H. and Miller, N. E. (1942). A multi-purpose learning demonstration apparatus. *J. exp. Psychol.*, **31**, 163–170.

Perin, C. T. (1942). Behavior Potentiality as a joint function of the amount of training and the degree of hunger at the time of extinction. *J. exp. Psychol.*, **30**, 93–113.

Saltzman, J., and Koch, S. (1948). The effect of low intensity of hunger on the behaviour mediated by a habit of maximum strength. *J. exp. Psychol.*, **38**, 347–370.

Skinner, B. F. (1938). *The Behaviour of Organisms*. New York.

Skinner, B. F. (1948). Are theories of learning necessary? *Psychol. Rev.*, **57**, 193–216.

Williams, S. B. (1938). Resistance to extinction as a function of the number of reinforcements. *J. exp. Psychol.*, **23**, 506–521.

THE INFLUENCE OF PREVIOUS MOVEMENT AND POSTURE ON SUBSEQUENT POSTURE

BY

C. V. JACKSON

(*From the Institute of Experimental Psychology, University of Oxford*)

An experiment with 21 subjects has confirmed the findings of other investigators that previous posture affects subsequent posture—the phenomenon of postural persistence. There seems also to be a tendency for the arm to be judged horizontal when in fact above the horizon, no matter which posture has been previously adopted—a "constant upwards effect." It has also been found that the direction of previous movement affects subsequent posture. After movement, an overshooting effect adds to a "constant upwards effect."

Various explanations of the phenomenon are discussed, and persistence is considered as an anomaly of postural recognition.

I

Introduction

The phenomenon of Postural Persistence has been described by Hoff and Schilder (1925), and studied in detail by Selling (1930). It consists in an unwitting deviation from, for example, the horizontal posture, when the subject attempts to bring his arm to the horizontal (in the absence of vision), after holding it in another posture for some time.

In the experiment of Hoff and Schilder the subject held his two arms horizontally in front of him, then raised one to 45° above the horizontal, keeping it there for about 30 seconds of "exposure." He then closed his eyes, and tried to bring the elevated arm to the same position as the other. In fact the arm remained somewhat raised.

Selling (1930) studied in 124 subjects the influences of duration of "exposure" time, of sex and of age and of which arm was tested for the phenomenon. He retained the two arm method, although Hoff and Schilder had pointed out that the effect could be obtained equally well if only one arm was used.

Dusser de Barenne (1934) expressed the opinion that the persistence was noticeably variable in any one subject, and did not occur in many individuals. He stated that voluntary activity could easily modify and obscure the phenomena, and it was therefore "not permissible to bring them into parallelism with the true brain-stem reflexes of Magnus."

Since postural persistence was to be studied in an experiment involving a conflict of visual and postural cues, it was decided to investigate the effect of previous posture on subsequent posture, in order to confirm Hoff and Schilder and Selling. It was also decided to investigate the effect of voluntary movement on the phenomenon.

II

Method

Small alloy clips were placed on the middle finger of each hand. These were shaped to form pointers which projected over the nail and about 5 mm. beyond the rounded end of the finger.

A point at the top of the right humerus was then chosen as the pivot point and proximal extremity of that arm. The determination was somewhat arbitrary, but sufficiently exact for the purpose. The lengths from this to the tip of the pointer (arm length), and to the floor (shoulder height) were then measured. It was assumed that these measurements were approximately the same on the left side.

The subject was then instructed: "You are to swing your arm five times completely round forward, and finish at the horizontal." This was demonstrated, and the subject told: "When you stop moving, you may correct the position of your arm as much as you like. This is not an exercise in correct stopping, but in judging the horizontal. Try to keep your fingers and arm straight all the time, and to come back to the same position at the end of every trial." The subject was asked to report any thoughts which might be relevant to the experiment. Ten trials with each arm were then undertaken, with the subject standing and blindfolded, in the sequence right, left, right, left, etc. (Action 1), the height of the pointer from the floor being noted after every judgment. The subject then rested for a short period and similar instructions were given to swing the arm backwards. The same procedure was followed (Action 2).

After a rest, the subject was told: "This time the procedure is somewhat different. You start with both arms folded. On the command 'Drop right,' you loosen your right arm and let it hang freely downwards. After five seconds, I will say 'Raise,' and you must bring the arm up to the horizontal." Having ascertained that the instructions had been understood, the experimenter continued, "We will try first the right, then the left as before. The arm which is not being tested in any particular trial may either be left folded, or also dropped to the side." Again ten trials of each arm were noted (Action 3).

After a rest, the subject was instructed: "This series is the same as the last except that, instead of dropping the arm freely downwards, you raise it to the vertical position. Thus, I will say 'Raise right,' and five seconds later '*Lower* to the horizontal.'" Again ten trials of each arm were undertaken (Action 4).

After another rest, the subject was instructed, "This time you also start with arms folded. When I say 'Extend Right,' stretch your arm forward in the horizontal position. After five seconds I will say 'Down and up.' You should then drop the arm freely downwards to the lowest position and *immediately* return it to the horizontal. Only the right arm will be tested." Ten trials were undertaken (Action 5). A further ten trials of the right arm only followed the instructions, "This series is exactly the same as the last except that, when your arm has been in the horizontal position for five seconds, I will say 'Up and down.' You should raise it to the vertical position and *immediately* return it to the horizontal." (Action 6).

A few subjects were also tested with trials of Actions 1 to 6 following each other in random order, the only sequence being left, right, left, but this procedure was very tedious and did not give rise to results different from those obtained by the other procedure.

III

Results

The results of 21 subjects were submitted to analysis of variance. The within-subject variance differed considerably from one subject to another; as might have been expected, some subjects were more consistent in their judgments than others.

In the first analysis, the results of all subjects ($N = 21$) in actions common to both arms were analysed. It was assumed that, when the marker placed on the finger tip was at shoulder-height (as defined above), the arm was horizontal. The height of the final judgment after deduction of the shoulder height was converted to the tangent of the angle which the arm made to the horizontal. In this way, differences of shoulder height and arm length from one individual to another were obviated. Since most deviations were upward from the horizontal, this direction was termed positive, and, for convenience, the tangents were multiplied by 1,000.

The means obtained in this way are shown in Table I below.

TABLE I

$N = 21$	*Mean deviations above horizontal (tangent units)*					
Actions	1	2	3	4	*Overall*	
Left Arm	11·43	40·92	−12·39	30·25	17·55	S.E. ±1·31
Right Arm	19·75	67·83	12·40	41·11	35·27	
Mean	15·59	54·37	0·01	35·68	26·41	S.E. ±0·93
	S.E. ±1·86					

S.E. of Entries in the Body of the Table = ±2·63.

The within-subject differences were so great, that before drawing conclusions from Table I, it was decided to test all effects against their interactions with subjects.

Following this, it was found that the effects left arm × right arm, and arms × actions were significant at only P = 0·05, and Action 1 × Action 2 and Action 3 × Action 4 at P = 0·01.

The barely significant arm × arm effect accords with Selling's results. He reported that more persistence was shown by the left arms of right-handed subjects. Of our 21 subjects only one man was left-handed, and two others were partly right-handed, probably shifted sinistrals. The remainder were fully right-handed.

Table I shows that, after exposure downwards (Action 3), the left arm judgments were below those of the right arm. On the other hand, after exposure upwards, judgments of the left arm were not above those of the right. In Action 3 only the left arm finished *below the horizontal.*

TABLE II

	Mean deviation above horizontal					*S.E. of individual means*
Actions	1	2	3	4	*Overall*	
Male N = 13	4·81	62·96	1·65	28·89	24·58 S.E. ±1·18	±2·36
Female N = 8	33·11	40·41	−2·67	46·69	29·39 S.E. ±1·15	±3·01

The overall means showed no significant sex differences (Table II) but there was a suggestion that the judgments of women were influenced less by previous movement than were those of men, and more by previous posture. After taking out the subjects, arms, and actions effects, there remained a residual of the total variance indicating variations between trials within runs of ten. This revealed no practice nor cumulative effects. As a further check against cumulative effects several subjects were tested with trials of the different series and arms fully randomized. No such effects were found.

In the second analysis, the results of all subjects (N = 21) in actions common to the *right arm only* were analysed, the means being obtained in the same way as in the first analysis. Table III shows that there were no overall sex effects and normal postural persistence is again apparent (Actions 3 and 4).

TABLE III

	Mean deviation above horizontal					*S.E. of individual means*
Action	3	4	5	6	*Overall*	
Male N = 13	12·48	35·28	66·25	37·31	37·83 S.E. ±1·54	±3·08
Female N = 8	12·26	50·58	75·61	36·85	43·82 S.E. ±1·96	±3·93
Means N = 21	12·40	41·11	69·82	37·13	40·11 S.E. ±1·21	
	S.E. ± 2·42					

It might be expected that "exposure" at a horizontal position (Actions 5 and 6) would result in a final judgment midway between those of Actions 3 and 4. The figures did not confirm this except in the case of women in Action 6.

The overshooting of the true value on raising to the horizontal in Action 5 is comparable to that in Action 2 (swinging backwards), except that men were more subject to the effect in Action 2, and women in Action 5.

The subjects × subjects, subjects × actions, actions × actions, and sex × actions interactions were all significant even when tested against the residual term (P = 0·01), and the actions effect remained significant (at P = 0·01) when tested against the subjects × actions effect.

TABLE IV

Question	*Answer*	
1. How is horizontal judged?	(*a*) Felt to be horizontal	8
	(*b*) Feeling—relation to shoulder ..	2
	(*c*) Feeling—relation to trunk	2
	(*d*) Feeling—undefined	1
	(*e*) By visual imagery	2
2. How are you sure that arm is horizontal?	(*a*) Felt to be horizontal	4
	(*b*) Feeling—relation to shoulder ..	2
	(*c*) By visual imagery	1
	(*d*) By returning to same position ..	1
	(*e*) By moving back same distance ..	1

In the third analysis, the results of subjects 11 to 21 only were analysed, the right arm alone being considered, and only the results of Actions 5 and 6. It will be remembered that both the initial (exposure) values of the horizontal and the

final judgments of these subjects were measured, since it was thought that the initial value might have influenced the final. This was found to be so. The regression of the final value on the initial was very highly significant (regression coefficient = 0·574; S.D. = 0·059). Even after adjustment for this regression, the effects of subjects and actions with their various interactions were still highly significant, but there was no sex effect before or after adjustment. Thus the different effects of the movements upwards and downwards from an approximately horizontal exposure were not only due to variations in this initial exposure position, but were also a result of the movements.

From questions asked after the experiment, it was apparent that most subjects judged the horizontal by "knowing how the arm feels when horizontal." Table IV summarizes the answers to the questions.

IV

Discussion

Selling's results were confirmed, in that it was shown that a previous posture lasting for as little as five seconds influenced subsequent posture.

Since most of our subjects were right-handed, and since the normal effect of postural persistence was somewhat more evident in the left arm than in the right (although in fact the differences were not significant at more than $P = 0{\cdot}05$), Selling's conclusions on the effect of handedness were on the whole confirmed.

However, in our Action 3 (arm exposed downwards), judgments of the horizontal by the right arm were *above* horizontal (Table I) which is a reverse tendency.

It is not clear whether Selling found this reverse tendency, since he mentioned no plus nor minus values. If he did not, then Table V compares his results with those of our experiment converted to angles.

TABLE V

5 secs. exposure	*Selling*	*Present experiment*
Ages	18–23	20–30
N	48	21
Right Up	+1·32 deg.	+2·33 deg.
Left Up	+2·02 deg.	+1·75 deg.
Right down ..	−1·27 deg.	+0·70 deg.
Left down.. ..	−1·88 deg.	−0·70 deg.

It will be noted that the age range of our subjects was greater, though their number was less. However, whereas our angles were calculated according to the arm length of every individual separately, those of Selling were obtained from the average lengths of his two sex groups. Also Selling's subjects exposed at 45° from the horizontal, and their arms were held in that position during exposure, whereas ours exposed at 90° with free arms.

Bearing in mind these differences in procedure, the two sets of results would not seem to disagree, with the exception of our anomalous right-down value.

In Actions 1 and 2, the arm was held in the lowest (vertically downwards) position and then swung five times. The swings lasted from 3 to 7 seconds with individual differences. One might therefore expect a reasonably constant posture, influenced by the initial downwards position, since postural persistence has been shown to persist up to 60 seconds (Eidelberg [1926] quoted by Selling p. 6). However, Tables I and II showed that in both actions the final judgment was *above* the horizontal for both arms and both sexes, and that the magnitude of the upwards deviation was greater in Action 2 (swinging backwards). Thus the direction of swinging had a marked influence on the final horizontal.

It might be supposed that the final scores of Actions 1 and 2 were the result of a "constant upwards effect" summing with the movement effect of overshooting. The data also agree with Sandstrom's (1951) conclusion that men tend to point below eye level, and women to eye level.

From the introspective evidence summarized in Table IV, it appears that most subjects determined the horizontal when blindfolded by "knowing how their arms feel when horizontal." The consistency with which individuals judged the horizontal varied considerably, but the subjective awareness of "how the arm feels when horizontal" was not much divergent from the objective posture (initial values Actions 5 and 6), unless the arm had previously been held in an "exposure" posture which differed markedly from the horizontal.

Since most experiments on Postural Persistence have centred round the horizontal posture, it might be thought that the phenomenon resulted from poor judgment of this arbitrary position. However, the effect of persistence can be shown, it seems, from any posture to any other posture. Also, if a mark is made anywhere up a wall within arm's reach, and the arm is hidden by a screen, it will be found that the arm can be brought opposite this mark fairly accurately, even though there is no sight of the arm to give visual "feed-back." If, however, the arm has been previously "exposed," the usual postural persistence effect ensues.

Thus our experiment has confirmed that previous posture influences subsequent posture, and has shown that movements of long duration influence subsequent posture in the direction of overshooting, while movements of short duration have a lesser effect, the previous posture again assuming importance.

In previous studies, three explanations of this phenomenon have been offered. It was thought to be due: (1) to "persistence of muscular tension and innervation" (presumably as a result of central after-discharge); (2) to "sensory adaptation"; or (3) to the characteristics of "a central apparatus."

The first explanation was rejected by Hoff and Schilder on the grounds that, if the "exposure" was below the horizontal, the final judgment was also below the horizontal, and that persistence is apparent even when the arm is relaxed. In addition, to maintain a horizontal arm posture apparently requires more "muscular tension and innervation" than to maintain other arm postures for an equal time. Also Hick (1953) found that similar phenomena were much affected by psychological factors, and concluded that they were not due solely to physiological after-discharge.

The second explanation was favoured by Hoff and Schilder (1925) and by Hick (1953), who studied aftercontractions of a similar nature and concluded that there must be a process counteracting sensory adaptation "otherwise a voluntary isometric contraction intended to be of constant force would steadily increase." The process postulated though "initially equal to that of sensory adaptation declines more rapidly in time," and "is the recovery from that which is felt as fatigue." He remarks that "sensory adaptation . . . is something we have to live with, and learn to allow for, as far as possible," and concludes that "information from less-adapted

or more quickly recovering sense organs would be available" to compensate partially for the effect of sensory adaptation.

Thus Hick has adopted the view that the phenomena result primarily from the properties of the nervous receptors involved, and has suggested how this may be reconciled with his finding that "after-contraction is . . . strongly affected by psychological factors" (1953, p. 169).

The third explanation, in terms of "a central apparatus" was adopted by Hoff and Schilder. The nature of this was not specified, but it was probably similar in conception to Head's (1911) Postural Schema, which has survived many attempts at clarification.

The present writer visualises a comparison (more or less conscious depending on the individual) being made between the total sensory input at the moment and the stored past impressions of "how the arm feels when horizontal" (Table IV). The latter aspect seems to deserve more emphasis than was given to it by Hick.

After "exposure," sensory adaptation (and perhaps fatigue) alter the total input stream of afferent impulses from the arm. The altered stream then "matches" with whatever neuronal processes underlie awareness of "how the arm feels when horizontal," when in fact the objective posture is incorrect.

In this way, postural persistence may be regarded as an *illusion* in the posturo-kinesthetic sense-modality, or as a normally occurring *anomaly* of postural recognition.

A more detailed explanatory hypothesis must evidently await advances in our understanding of the operation of the many different types of sensory receptors involved, of the neurophysiology of fatigue, and of awareness of posture generally.

The author is grateful to Professor G. Humphrey for permission to carry out the work, and to Dr. M. R. Sampford for statistical help. He was working with a grant from the Medical Research Council,

References

Dusser de Barenne, J. G. (1934). The Labyrinthine and Postural Mechanisms. Ch. 4 in *Handbook of General and Experimental Psychology* ed. Murchison. Worcester, Mass.

Eidelburg, L. (1926). Quantitative Untersuchungen der Lagebeharrung. *Mschr. Psychiat. Neurol.*, **61**, 101.

Head, H., and Holmes, G. (1911). Sensory Disturbances from Cerebral Lesions. *Brain*, **34**, 102–190.

Hick, W. E. (1953). Some features of the After-Contraction Phenomenon. *Quart. J. exp. Psychol.*, **5**, 166–170.

Hoff, H., and Schilder, P. (1925). Uber Lagebeharrung. *Mschr. Psychiat. Neurol.*, **58**, 257.

Sandstrom, C. I. (1951). *Orientation in the Present Space.* Almqvist and Wiksell. Stockholm.

Selling, L. S. (1930). An Experimental Investigation of the Phenomenon of Postural Persistence. *Arch. Psychol., No.* 118, *New York.*

AN EXPERIMENT WITH OCCASIONAL FALSE INFORMATION

BY

J. A. LEONARD

(*Nuffield Research Unit into Problems of Ageing, The Psychological Laboratory, Cambridge*)

An experiment is described in which subjects were presented, under one condition, with advance information of doubtful value. The evidence suggests that advance information was accepted, although the uncertainty created had a detrimental effect on the subjects' performance.

I

Introduction

In an earlier paper on advance information in sensori-motor skills (Leonard 1953) "receptor-effector" anticipation was defined and its role discussed. The present paper is a sequel.

The work was carried out with the stylus apparatus described in the earlier paper. With this apparatus subjects could be given information either at or before a common action point. It consisted of a five light display and a control panel with five discs flush with the plywood forming a pentagon corresponding to the display. There was one disc in the centre of the control panel, and subjects had to slide a stylus between this centre disc and an outer one according to the state of the display. In condition (1) the display changed when the centre disc was touched, the indicated disc having been touched first; in condition (2) the display changed when the indicated disc was touched, the centre disc having been touched first. Thus for each response both indicated disc and centre disc had to be touched, but whilst in the first condition subjects could not know what to do beyond the centre disc until they actually got there, this information was provided in advance, i.e. at the indicated disc, in the second condition. In this condition therefore subjects could know all the time what to do next; on leaving the centre disc for the then indicated disc they knew that they would have to return to the centre disc and on reaching the indicated disc they were given the information concerning activity beyond the centre disc. There was thus something like "continued certainty." It was pointed out in the preceding paper that in the condition where information was given only at the centre disc, the point of action as it were, there had to be something of a hiatus in the flow of definite information. But under both conditions used, whenver information was presented, it could always be relied upon in the sense that a disc once indicated would remain "correct" until touched. This is of course an assumption which always has to be made whenever a response is initiated.

It has been known for some time that with simple responses there is a period of about 0·12 seconds after the initiation of a response during which the response can no longer be checked. If the display changes during this period, we may be aware of this change, but we can do nothing to arrest the movement (Guenther, 1911, Bates and Hugh Jones, 1944). The results of the stylus experiment reported in the previous paper suggested that subjects in condition (2) had initiated the appropriate activity for the stage beyond the centre-disc before reaching the centre disc. Pushing the argument a little further, we might say that they had accepted the advance information and were acting upon it in the firm "knowledge" that their action would still be appropriate. It seemed interesting to find out what would happen when the conditions no longer provided an adequate basis for firm knowledge, i.e. when the

display was, in fact, liable to change again when subjects got to the centre disc and the new and not the old display would have to be responded to. The question now was: if information given ahead of the point of action has a measure of uncertainty attached to it, will it still be accepted?

II

Methods and Procedure

Apparatus

The stylus apparatus has been described in the earlier paper (Leonard 1953). In the present experiment sequences of 50 stimuli only were used. An additional bank on the uniselector was used to provide occasional contacts in parallel with existing stimulus positions. When a subject touched the centre disc at a time such a contact was reached in the series, a relay changed the display by switching over from the currently used bank to yet another bank. The relay would be held until the next correct response returned it to its resting state and the main bank came back into use. The second display change therefore was not produced in the same manner by the apparatus as in conditions (1) and (2). The apparatus latency can reasonably be assumed to have been very slightly shorter, if anything, in the former case. It was necessary in this experiment to introduce an additional auditory error indicator in the form of a single stroke on a bell. Pilot runs had shown that subjects otherwise might not be aware at all of the second display change before they had reached the disc originally indicated but now incorrect, their indication of correctness of response being in fact that the display had changed. Usually one experience of the bell per trial was enough to draw attention to the risk involved.

Conditions

The experiment was given under three conditions: the first two have already been described in the introduction. In condition 2a the display would *always* change when the subjects touched the indicated disc, having first touched the centre disc, but *sometimes* (see Table II) it would change again when the subject touched the centre disc. These occasional changes were spaced at random through the trial, and were never entirely in the same order for successive trials.

Procedure

Table I shows the experimental plan. The two groups of three subjects each were not matched.

TABLE I

Experimental Plan

Day		1								2					
	Trial	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Group I	Condition	1	2	1	2	1	2	1	2	2a	2a	2a	2a	1	2
Group II		2	1	2	1	2	1	2	1	2a	2a	2a	2a	2	1

In all, subjects were given fourteen trials of 50 responses each on two successive days. They had the control-display relationship explained to them, were asked never to lift the stylus off the panel, and were told that their task was to make 50 correct responses in as little time as possible, making as few errors as possible. They were told that the apparatus would cut off when they had completed a trial. Possible errors were demonstrated and subjects were allowed to make a few practice moves before the first trial (mainly to show the experimenter that the instructions had been understood). They were not given any specific instructions on how to do the task. The subjects were told after each trial how much time they had taken. Apart from these general instructions there were specific instructions for the three conditions, given to each subject before he

was first exposed to them: (1) "Starting from the centre slide the stylus to the indicated disc and come back when one of the other four lights will come on. Only one lamp will be on at any one time." (2) "Starting from the centre disc slide the stylus to the indicated disc when one of the other four lights will come on. You must touch the centre disc before going on to the next indicated disc. Only one lamp will be on at any one time." Before their first experience of condition (2a) subjects were told that this was similar to condition 2 in that the display would always change when the indicated disc was touched, but that sometimes the display would change again when they touched the centre disc.

They were told that if the display changed again they would have to touch the now indicated disc. They did not know how many changes there would be, nor when they would occur in any one trial.

From a technical point of view it would have been possible to effect display changes at the centre on 28 occasions. A smaller number (see Table II) was eventually chosen because this enabled the experimenter to change the positions of these changes between trials.

It seemed advisable to vary the number as well as the position within the sequence at which secondary display changes could occur, in order to prevent subjects from learning to expect them.

Scoring was carried out in terms of total time, for all trials, to the nearest ½ sec.; and times for two components making up one response. (It will be appreciated that particularly under condition (2a) performance tended to be rather erratic so that the more detailed recording possible in the experiment reported earlier was no longer meaningful). The two components scored were (1) the time between touching the indicated disc and reaching the centre disc (EP & In), and (2) the time between touching the centre disc and reaching the indicated disc (C & Out). Those C & Out components of condition (2a), at the beginning of which there was no display change, were called "*unaffected*"; those at the beginning of which the display did change were called "*affected.*" Readings were taken to the nearest 0·025 sec. and only the 12th, 13th and 14th trials of each subject were scored in detail.

III

Results

TABLE II

Mean Times in Seconds per Trial for 6 Subjects

Trials	7/8	9	10	11	12	13/14
Condition						
1	46·2					44·9
2	38·6					36·2
2a		53·9	49·2	46·4	48·0	
Additional Display Changes		(17)	(17)	(12)	(16)	

Table II shows the composite final training scores for the 7th and 8th trials of conditions (1) and (2), the four trials under condition (2a), and the composite scores for the 13th and 14th trials under conditions (1) and (2). It should be noted that there was a marked slowing up on the first trial under condition (2a), the group settling down eventually to something near the level of performance under condition (1). The dip on the 11th trial was presumably due to the fact that there were fewer occasional changes in this trial than in the others.

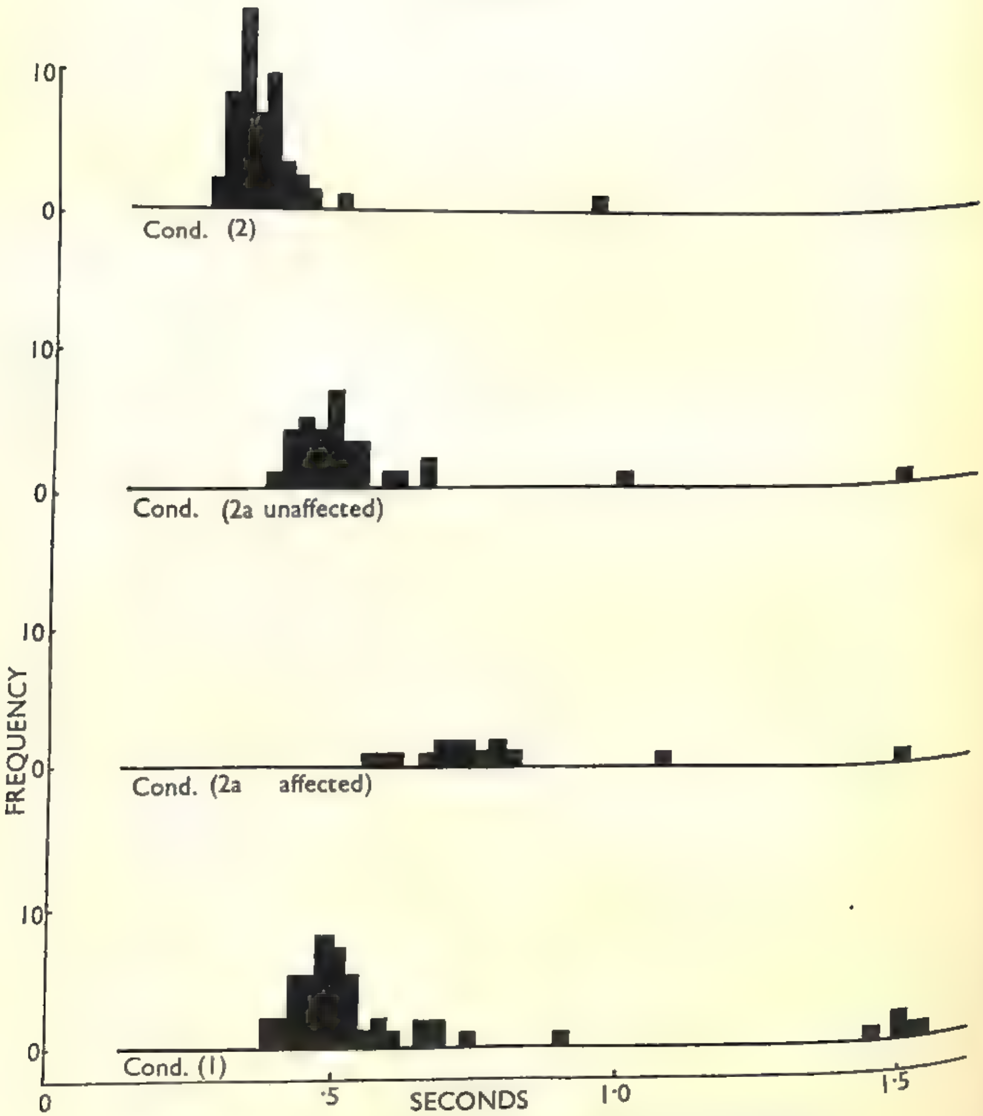
Component times were obtained for the 12th, 13th and 14th trials of each subject. The group means shown in Table III were obtained after excluding some very large readings which would have exaggerated the trends found. There were about 2 or

TABLE III

MEAN TIME IN SECONDS FOR THE TWO COMPONENTS PER RESPONSE IN TRIALS 12, 13/14

Condition		*C & Out*	*EP & In*
1		0·57	0·27
2		0·39	0·28
2a	unaffected	0·51	0·29
	affected	0·73	

FIGURE 1.



(C & Out) Distributions of one subject in main experiment

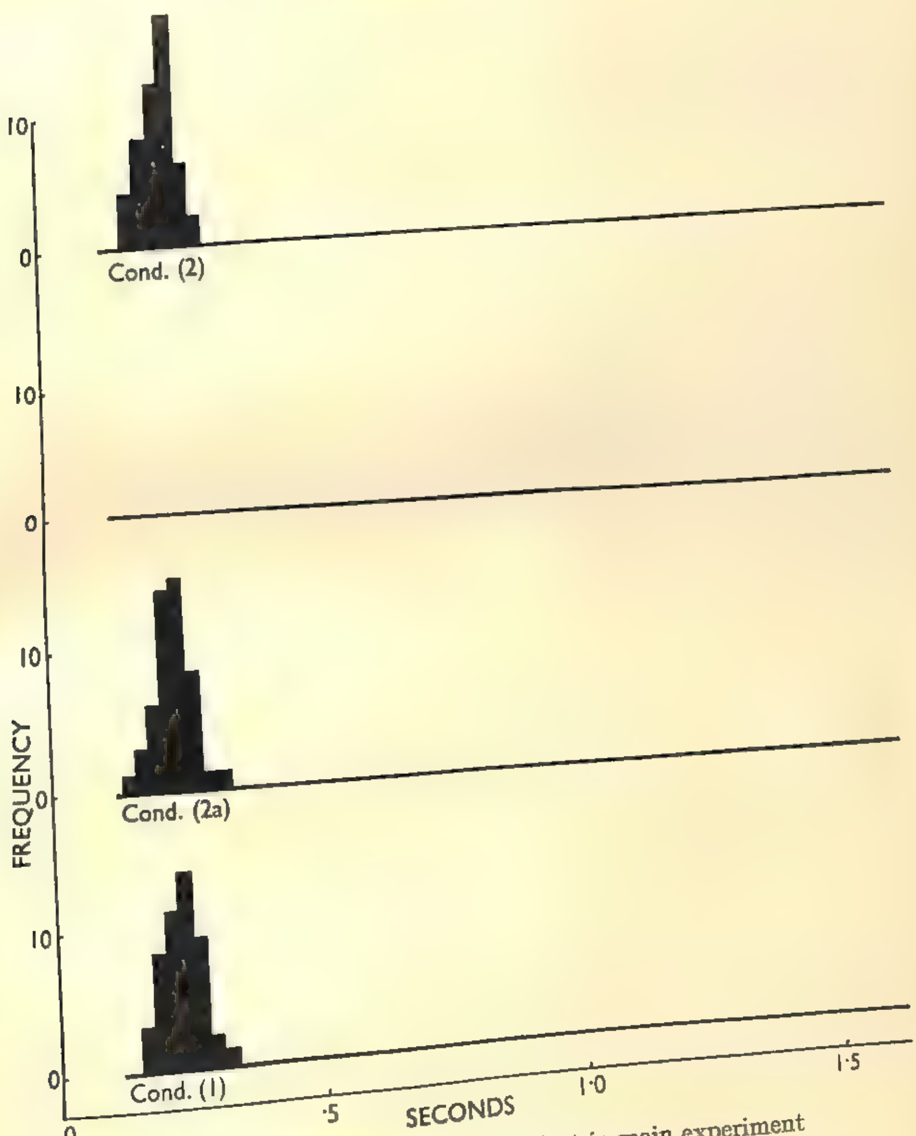
3 of these readings per distribution. A transformation to logarithms did not seem indicated.

The relationship between the three EP & In group means was not reflected in individual subjects' means and no consistent trend was noted in the three EP & In distributions for each subject. In the discussion below the C & Out components will be considered, neglecting differences between the other component.

The relationship between the group means obtained for the four C & Out components was reflected in all subjects. Similarly, the six sets of distribution showed the same general pattern.

A typical set of C & Out and EP & In distributions for one subject is shown on Figures 1 and 1a.

FIGURE 1a.



(EP & In) Distributions for one subject in main experiment

IV

Discussion

By instruction, subjects had to make 50 correct responses per trial in as little time as possible. In condition (2a) where the display changed again at the centre, they had at least two ways of complying with the instructions: They could either respond to the display only when they reached the centre disc; for the light which was on then would always be "correct." In other words they could respond as in condition (1). It was thought that subjects might eventually adopt this method, particularly since the occasional changes of display at the centre tended to impose a motor pattern more like that found in condition (1). Alternatively they could respond when reaching the indicated disc although on some occasions the light to which they had responded might turn out to be not correct. That is to say they could perform as in condition (2) risking however occasionally having to modify responses already initiated.

Considering the effects on the C & Out components one may compare the two expectations and the observed results in the following manner:

Expectation I	affected (2a) $=$ non-affected (2a) $=$ (1)
Result	affected (2a) $>$ non-affected (2a) $\approx$ (1)
Expectation II	affected (2a) $>$ non-affected (2a) $=$ (2)
Result	affected (2a) $>$ non-affected (2a) $>$ (2) $\approx$ (1)

One is therefore faced with the somewhat confusing situation that both expectations are partially fulfilled. This leads to an apparent contradiction. The fact that the non-affected components in condition (2a), i.e. when the display had not changed again at the centre, are similar to those in condition (1) suggests that information was accepted only at the centre disc in condition (2a). In this case a display change at the centre should however not have had any effect on the affected components. The fact that the affected components' distributions were always clearly separated from the non-affected components' distributions suggests that the display-change at the centre did have an effect. This in turn suggests that subjects had accepted information when it was presented at the indicated disc and had to re-organize their response when the display changed again at the centre.

But if in condition (2a) subjects did accept information at the indicated disc, why were the values for the non-affected components of that condition comparable to those of condition (1) rather than of condition (2)?

There seem to be two possible answers, and one cannot be certain whether these may not really be identical. It may be that subjects accepted information at the indicated disc but required confirmation at the centre-disc. Or it may be that the higher value of the non-affected components in condition (2a) reflects the need to watch the display more carefully and continuously in this condition. Both these answers suggest that the introduction of occasional additional display changes created a greater over-all uncertainty. If this contention is correct the numerical similarity between the values for the unaffected components of condition (2a) and those of condition (1) may be spurious, i.e. the values may reflect different operations.

There were therefore two effects which made performance under condition (2a) slower than under condition (2): a direct effect resulting from the actual additional display changes; and an indirect effect resulting from the greater over-all uncertainty.

On the whole the experimenter's observation of subjects' performance tended to be in line with the data obtained from the records. Subjects were clearly trying to maintain in condition (2a) the smoothness possible in condition (2). Far from opting to stop at the centre for every response they seemed to learn how to modify responses on the way to the previously indicated disc.

After the experiment it was suggested to each subject that it might perhaps have been more reasonable to perform in condition (2a) as in condition (1) and thus avoid the need to modify initiated responses. They did not think this a good idea. Neither did others to whom the experiment was demonstrated and who tried it out. It was possible deliberately to adopt the method of conditon (1) but it was not easy, and certainly not the spontaneous method adopted.

Subsidiary experiments carried out with different proportions of additional display changes, or subjects having condition (2a) without previous epxerience of the other two conditions bore out the main findings presented here.

The question asked in the introduction was: if the information given ahead of the point of action has itself a measure of uncertainty attached to it, will it be accepted? The evidence suggests that it was accepted, but that the efficiency with which it was used was reduced. This reduction in efficiency may reflect the greater uncertainty inherent in the task. It could also be due to the need to pay more attention to the display. In this particular setting—where of course making an error did not have very drastic consequences—subjects did seem to accept advance information even when it was of doubtful value.

References

BATES, J. A. V., and HUGH JONES, P. (1944). Physiological factors in the design of firing mechanisms. *M.P.R.C. of Med. Res. Coun.* B.P.S. 44/313. Feb., 1944.

GUENTHER. (1911). Reaktionsversuche bei Durch angsbeobachtungen. *Psychol. Stud.*, 7, p. 229 onw.

LEONARD, J. A. (1953). Advancei nformation in sensorimotor skills. *Quart. J. exp. Psychol.*, **5**, 141–149.

THE PERCEPTION OF SHORT PHRASES PRESENTED SIMULTANEOUSLY FOR VISUAL & AUDITORY RECEPTION

BY

G. H. MOWBRAY

From the Psychological Laboratory, The Johns Hopkins University, Baltimore, Maryland

Subjects were provided with outline maps that were incomplete in several details. Brief, simultaneous, visual and auditory instructions were given for completing some of the missing details. Certain items could be completed on the basis of direct information contained in one or other of the sensory modalities. Others, however, could be completed only because of their relation to details capable of location by direct instruction. Information important for the completion of map details was distributed randomly among short passages of unconnected words. All relevant visual and aural clues were presented simultaneously in every case. Opportunities for alternations of attention were curtailed.

Thirty-six subjects were randomly assigned to three experimental conditions, and to two groups that were given different instructions. One group was told that relevant information would always appear simultaneously, while the other group was not allowed this information.

The number of successfully located simultaneous pairs of items presented for direct location was found to be no greater than could be expected by chance. The total number of correctly located items was less than 50 per cent of the possible items. There was no difference in the number of correctly located simultaneous pairs of items between the "instructed" and the "uninstructed" groups. The "uninstructed" group did not learn in the course of the experiment that all relevant material was presented simultaneously. Significantly more correct completions were made with the visual material than with the auditory. It is concluded that successful division of attention did not occur.

I

Introduction

Recently, Hebb (1949), in proposing a new orientation to the problem of the organization of behaviour, touched briefly on the age-old argument concerning the division of attention. He adopted the position that attention may often be of a multiple nature since there is little, if any, evidence to support the commonly held view that attention to two things that are not organizable into a single unit is impossible. However, recently Mowbray (1953) has provided some data that, while not completely conclusive, do argue in favour of the unitary nature of attention.

Two main criticisms may be inveighed against this research. The method of visual presentation, while curtailing to a considerable extent the opportunities for alternations of attention, did not entirely eradicate them. The second objection concerns the nature of the experimental material. Prose selections were used, and they were of considerable length (154–220 words) so that the performance decrement exhibited from non-simultaneous to simultaneous presentation may have been as much a function of forgetting, owing to some interference between the two streams of conflicting stimuli, as it was an initial failure to perceive.

Purpose of experiment. Therefore, it was considered advisable to construct an experiment that would make use of brief verbal messages, with different auditory and visual messages presented simultaneously, and would in so far as possible prevent rapid alternation between sensory modalities. Accordingly, an experiment was designed that made use of brief verbal messages leading to a simple map-completion

test. The verbal redundancy of the instructions was reduced to a minimum, and their brevity served to keep them within normal memory span.

II

Experimental Method

Material. Seven coloured outline maps were traced on sheets of translucent drawing parchment about 22 in. long and 14 in. wide. The details on each map consisted of lakes (blue), towns (cross-hatched red lines), parks (green), highways (red), and groups of buildings (purple squares and rectangles). Very near to each detail a small rectangular cut-out was made in the maps as though a name had been deleted. Transparent plastic sheets of the same size were cut, and small rectangular holes corresponding in size and position to the cut-outs in the maps were made. The plastic sheets served to cover the maps and protect them from accidental marking and tearing.

Visual instructions for the location of the details on the maps were typed on a roll of recording paper 6 in. wide. A thyratron-driven motor pulled the tape past a viewing window cut in a grey-coloured, metal screen. The width of the viewing window was the same as the width of one line of type and its length enough to allow vision of the entire width of the recording tape. The lines of type were horizontal, but the viewing slit was slightly tilted down at the right. Thus only one of the four widely separated, but evenly spaced, words on each line could be seen at a time.

All names given to map details were one of ten colours (rose, blue, black, red, yellow, green, orange, purple, brown, or aqua), and no colour was used more than once for any detail on a given map. Information for the location of a detail was given in terms of alphabetical and numerical map co-ordinates. Thus, for instance, one of the towns on a map might be located from the following instructions: "Yellowtown at A-Seventeen."

There were six, eight, or ten details to be located on any one map depending on the condition being used in the test. The three test conditions utilized were as follows:

Condition I. There were six items to be located. Four of these could be located on the basis of direct information, i.e. they were associated with map co-ordinates, while the other two could be located only indirectly with reference to the location of the direct items. No co-ordinates were provided for these latter. The four details about which co-ordinate information was supplied consisted of two pairs of simultaneously presented items. Of each pair, one item was given visually and the other aurally.

Condition II. There were eight items to be located. Six of these could be located from information given in the form of map co-ordinates. The other two could be located indirectly as in Condition I. The six details about which co-ordinate information was supplied consisted of three pairs of simultaneously presented items. Each pair consisted of one visual and one auditory presentation.

Condition III. There were ten items to be located. Eight of these could be located from information given in the form of map co-ordinates. The other two could be located indirectly as in Conditions I and II. The eight details about which co-ordinate information was supplied consisted of four pairs of simultaneously presented items. As above, each pair consisted of one visual and one auditory presentation.

For every item that had to be located, there were on the map at least four possible correct answers considered from the standpoint of chance. That is to say, if there was one lake to be located, there was a total of four lakes on the map that could be the correct one. Likewise, if there were two lakes to be located, then there was a total of eight lakes on the map.

All of the information necessary for completing the required parts of a map was presented in four sections, but only a small proportion of any of the four sections was relevant to the task at hand. The remainder was unconnected words, similar in some ways to the key phrases, but making no sense overall. The length of each section ranged randomly from 24 to 44 words, and the relevant information was positioned by chance within a given section. A single section never contained more than two relevant items. Sometimes only one item was relevant, and for some of the conditions, entire sections contained no relevant information at all.

The auditory material was recorded on a magnetic tape recorder by a female voice, and was organized identically as described above. Overall synchronization of visual and auditory presentations was accomplished by careful timing and recording so that there was very nearly a word-for-word correspondence between the two channels. In all

sections the relevant visual and auditory phrases coincided temporally, and were of exactly the same length.

The speed of visual and aural presentation was approximately 106 words per minute. For normal reading this is an extremely slow speed, but it will be remembered that there were but four words to a 6-in. line necessitating unusually large eye-movements.

Procedure. The subjects were seated facing the visual pacer, while to their left was the auditory tape recorder, and to their right was a map-board holding the experimental map. They were given a sheet of typed instructions to read, and were told to take all the time they needed to understand them.

Following the reading of these very detailed directions, a few further instructions were given to emphasize some of the more important aspects of the experiment. The subjects were informed that the information would be given in four sections, but they were not told how many relevant items they were to expect in any given portion. They were told that some sections might not contain anything of use to them in their completions. They were then instructed that the auditory and visual displays would be halted at the end of each of the four sections to allow them to complete whatever details they could. They were asked not to attempt completions before the end of a section. Paper and pencil were provided to allow them to make any notations they desired during or after the presentation of any section. They were encouraged to guess at the locations if they had any reasonable basis for doing so, but were allowed only one guess for any one detail.

The subjects were given seven cards—one for each of the two practice maps and the five experimental maps—on which were printed the names of all of the items to be located on each map. They were allowed to keep the appropriate card available for reference throughout the entire experiment. Each map was an entity in itself and completely independent of every other.

Before beginning the experiment, one-half of all subjects were further instructed that all relevant pieces of visual and auditory information would appear simultaneously, and that there would *never* be any useful information in one channel without there also being at the same time some equally useful but different information in the other. This group was termed the "instructed" group. The other half of the subjects were not informed of this detail and hence were termed the "uninstructed" group.

Following this lengthy and involved briefing, the two practice maps were completed, during which time all misunderstandings and misapprehensions were cleared up. Responses were made, through cut-outs in the maps and covering plastic sheets, on to paper inserted under the maps.

The experiment was completed in a single session requiring about 1 hr. and 15 min. in all—including instructions, practice, and test maps.

At the end of the experiment all subjects in the "uninstructed" group were asked the following question: "What can you tell me about the presentation of the relevant items of information that appeared in the visual and auditory series?" If they did not reply that the relevant bits had appeared simultaneously in both channels, they were then asked specifically if they knew that such had been the case in every instance, and their replies were recorded.

Subjects. A total of 36 subjects was assigned at random to one or other of the three conditions and to the "instructed" or the "uninstructed" group. They were university undergraduates, graduate students, scientific research workers, and a few others participating to a large extent in university life. They ranged in age from 19 to 36 years.

III

Results and Discussion

In considering the results, the total emphasis will be placed on those items about which direct information was supplied. Those items that could be located only with reference to the position of directly given ones were originally included in the experiment to provide supplementary information in case simultaneous reception proved possible. As will be shown subsequently, their only function was to provide two additional guesses per map in the experimental situation.

Simultaneous pairs. Since the major purpose of this experiment was to determine whether simultaneous auditory and visual receipt of information could occur, the

first analysis will be concerned with the number of times that both the visual and auditory items of a simultaneous pair were correctly perceived. If simultaneous reception of information is impossible, no pairs should have been perceived except as the result of chance guessing.

Table I shows the number of successful completions of both of a pair of items presented simultaneously for direct location. The results are tabulated for the three

TABLE I

SUCCESSFUL COMPLETIONS OF BOTH OF A PAIR OF ITEMS PRESENTED SIMULTANEOUSLY FOR DIRECT LOCATION

	Experimental condition			
	I	*II*	*III*	*Total*
"Instructed" group				
Number	5	17	12	34
Possible	60	90	120	270
Per cent.	8·3	18·9	10·0	12·2
"Uninstructed" group				
Number	9	8	12	29
Possible	60	90	120	270
Per cent.	15·0	8·9	10·0	10·7

conditions and for the two groups ("instructed" *vs.* "uninstructed"). The extreme right-hand column gives the total pairs summed over the conditions. An analysis of variance was performed on the data from which this Table was compiled. The results are given in Table II where it can be seen that neither the instruction given to subjects nor the experimental conditions contributed any significant variance.

TABLE II

ANALYSIS OF VARIANCE OF THE PER CENT. CORRECT COMPLETIONS OF BOTH OF TWO ITEMS PRESENTED SIMULTANEOUSLY FOR DIRECT LOCATION

Source	*Sum of squares*	df	*Variance estimate*	F
Instruction ..	10·03	1	10·03	—
Conditions ..	71·17	2	35·58	—
Interaction ..	413·38	2	206·69	1·9
Within cells ..	4206·17	30	140·21	
Total	4700·75	35		

From Table I it can be seen that some simultaneous pairs were apparently successfully located. Although their total number was small, the important thing to be determined is whether or not they were appreciably more numerous than might be expected on a chance basis. For the purpose of analysis, unfortunately, exact computation of chance probabilities is impossible since the correct location of one or more items considerably altered the odds of further guessing of correct locations.

Even though exact determination of probabilities for the successful location of two simultaneously presented pairs of items was not possible, it was possible to

calculate some approximations. This was done for all three conditions. In every case the probabilities were figured for the circumstance that provided the most unfavourable opportunities for correct guessing. As an example, for Condition I it was assumed that one-half of the directly given items were located correctly (e.g. all of the visual or all of the auditory). In this case, that accounted for two of the six items to be located. Since only the direct items are being considered here, the subject had four guesses to locate correctly the remaining two locations. He knew which of the five categories of items (towns, lakes, parks, buildings, and highways) were to be guessed from, and how many of each category were to be located. This information was given him before the start of the experiment, it will be remembered. Assuming further that only three of the five possible categories were used, it was possible to ascertain the average number of simultaneous pairs that would be expected to be located successfully for a given map. This theoretical frequency would be the lowest that could be expected if all of the maps and all of conditions were as unfavourable as possible for making correct locations by guess alone. It is perhaps unnecessary to point out that many of the maps and conditions provided probabilities considerably greater than those considered in this type of evaluation.

Using this method of analysis, it was determined that for Condition I an average of 0·34 correct simultaneous locations could be expected per map. For Conditions II and III, the figures were 0·47 and 0·51, respectively. Since there was a total of 60 maps for each condition (both groups), the expected numbers of correctly located simultaneous pairs would be 20·4, 28·2, and 30·6 for the three conditions. A look at Table I shows that the obtained figures were 14, 25, and 24. Therefore, it seems fair to conclude that subjects did no better than would have been expected from random guessing.

Effect of instructions. The fact that no more successful completions of simultaneously presented items were made by the "instructed" group than by the "uninstructed" one is subject to two interpretations. Either the "uninstructed" subjects became aware of the simultaneous presentation in the practice trials or early in the experimental trials, or else, under conditions that provided no opportunities for rapid shifts of attention, only one perception at a time was possible, and the instructions were of no assistance. That the latter interpretation is the correct one is illustrated by the answers to the questions put to the "uninstructed" group at the end of the experiment. Of the 18 subjects in this category, not one volunteered the correct answer when asked what they could tell the experimenter about the presentation of the relevant items of information that had appeared in the visual and auditory series. When asked directly if they knew that the presentation had always been simultaneous, 2 replied that they had noticed it happening once or twice, whereas the other 16 denied ever having noticed it. Many were vehement in their denials and expressed outright surprise, and in some instances, sheer disbelief that such had really been the case.

In the light of this result, it seems fair to state that the reason for the lack of difference between the two groups in the above analysis resides in the inability of the "instructed" group to take advantage of their knowledge.

Total items. Another way of considering the data from the items about which direct instructions for location were given is to consider the total percentage of such items correctly located. Table III presents this information without regard to the two groups of subjects. Here it can be seen that in no case were more than 50 per cent. of the items presented correctly located. If for single-modality presentation, the task proved so difficult that only a relatively small percentage of items could be located anyhow, this result would not prove very meaningful. However, the speeds

of stimulus presentation were very slow (106 words per minute) and the single modality task was well within the capabilities of a normal observer. Evidence of this is provided by some of the individual results. One subject correctly located all of the visual presentations and none of the aural ones, whereas several others while locating only one or two of the visual or aural presentations correctly located nearly

TABLE III

PERCENTAGE OF DIRECT ITEMS CORRECTLY LOCATED
(Visual and Auditory)

Condition I	*Condition II*	*Condition III*
41·7	45·8	44·6

all of the items presented in the other modality. It seems safe to say, then, that not only is there no evidence to show that simultaneous perceptions were accomplished in this situation, but, on the contrary, there is some possibility that actual interference was produced by the simultaneous presentations.

Modality differences. The number of correct responses mediated by the visual and the auditory channels for the directly presented items provides some interesting results. These data are summarized in Table IV. Here it can be seen that the visual

TABLE IV

TOTAL NUMBER OF CORRECT COMPLETIONS OF DIRECTLY GIVEN VISUAL AND AUDITORY ITEMS FOR BOTH EXPERIMENTAL GROUPS AND FOR THREE CONDITIONS OF INFORMATION

Condition	*Group*	*Visual*	*Auditory*
I	Instructed	33	13
	Uninstructed	38	16
II	Instructed	50	42
	Uninstructed	41	32
III	Instructed	53	49
	Uninstructed	81	31
Total		296	183

completions are consistently greater in number than the auditory ones for both groups and for the three conditions utilized in the experiment. An overall test for the significance of the difference between the visual and auditory scores considered without respect to groups or conditions yielded a critical ratio of 3·1 which indicates significance ($p < 0{\cdot}002$).

These latter results are in line with those reported earlier by Mowbray (1953) relating to the dominance of the visual channel over the auditory with conflicting simultaneous combinations. The explanation for this visual dominance presumably resides in the nature of the auditory and visual presentation. The constantly moving visual display required an active effort on the part of the subject to the apparent

calculate some approximations. This was done for all three conditions. In every case the probabilities were figured for the circumstance that provided the most unfavourable opportunities for correct guessing. As an example, for Condition I it was assumed that one-half of the directly given items were located correctly (e.g. all of the visual or all of the auditory). In this case, that accounted for two of the six items to be located. Since only the direct items are being considered here, the subject had four guesses to locate correctly the remaining two locations. He knew which of the five categories of items (towns, lakes, parks, buildings, and highways) were to be guessed from, and how many of each category were to be located. This information was given him before the start of the experiment, it will be remembered. Assuming further that only three of the five possible categories were used, it was possible to ascertain the average number of simultaneous pairs that would be expected to be located successfully for a given map. This theoretical frequency would be the lowest that could be expected if all of the maps and all of conditions were as unfavourable as possible for making correct locations by guess alone. It is perhaps unnecessary to point out that many of the maps and conditions provided probabilities considerably greater than those considered in this type of evaluation.

Using this method of analysis, it was determined that for Condition I an average of 0·34 correct simultaneous locations could be expected per map. For Conditions II and III, the figures were 0·47 and 0·51, respectively. Since there was a total of 60 maps for each condition (both groups), the expected numbers of correctly located simultaneous pairs would be 20·4, 28·2, and 30·6 for the three conditions. A look at Table I shows that the obtained figures were 14, 25, and 24. Therefore, it seems fair to conclude that subjects did no better than would have been expected from random guessing.

Effect of instructions. The fact that no more successful completions of simultaneously presented items were made by the "instructed" group than by the "uninstructed" one is subject to two interpretations. Either the "uninstructed" subjects became aware of the simultaneous presentation in the practice trials or early in the experimental trials, or else, under conditions that provided no opportunities for rapid shifts of attention, only one perception at a time was possible, and the instructions were of no assistance. That the latter interpretation is the correct one is illustrated by the answers to the questions put to the "uninstructed" group at the end of the experiment. Of the 18 subjects in this category, not one volunteered the correct answer when asked what they could tell the experimenter about the presentation of the relevant items of information that had appeared in the visual and auditory series. When asked directly if they knew that the presentation had always been simultaneous, 2 replied that they had noticed it happening once or twice, whereas the other 16 denied ever having noticed it. Many were vehement in their denials and expressed outright surprise, and in some instances, sheer disbelief that such had really been the case.

In the light of this result, it seems fair to state that the reason for the lack of difference between the two groups in the above analysis resides in the inability of the "instructed" group to take advantage of their knowledge.

Total items. Another way of considering the data from the items about which direct instructions for location were given is to consider the total percentage of such items correctly located. Table III presents this information without regard to the two groups of subjects. Here it can be seen that in no case were more than 50 per cent. of the items presented correctly located. If for single-modality presentation, the task proved so difficult that only a relatively small percentage of items could be located anyhow, this result would not prove very meaningful. However, the speeds

of stimulus presentation were very slow (106 words per minute) and the single modality task was well within the capabilities of a normal observer. Evidence of this is provided by some of the individual results. One subject correctly located all of the visual presentations and none of the aural ones, whereas several others while locating only one or two of the visual or aural presentations correctly located nearly

TABLE III

PERCENTAGE OF DIRECT ITEMS CORRECTLY LOCATED
(Visual and Auditory)

Condition I	*Condition II*	*Condition III*
41·7	45·8	44·6

all of the items presented in the other modality. It seems safe to say, then, that not only is there no evidence to show that simultaneous perceptions were accomplished in this situation, but, on the contrary, there is some possibility that actual interference was produced by the simultaneous presentations.

Modality differences. The number of correct responses mediated by the visual and the auditory channels for the directly presented items provides some interesting results. These data are summarized in Table IV. Here it can be seen that the visual

TABLE IV

TOTAL NUMBER OF CORRECT COMPLETIONS OF DIRECTLY GIVEN VISUAL AND AUDITORY ITEMS FOR BOTH EXPERIMENTAL GROUPS AND FOR THREE CONDITIONS OF INFORMATION

Condition	*Group*	*Visual*	*Auditory*
I	Instructed	33	13
	Uninstructed	38	16
II	Instructed	50	42
	Uninstructed	41	32
III	Instructed	53	49
	Uninstructed	81	31
Total		296	183

completions are consistently greater in number than the auditory ones for both groups and for the three conditions utilized in the experiment. An overall test for the significance of the difference between the visual and auditory scores considered without respect to groups or conditions yielded a critical ratio of 3·1 which indicates significance ($p < 0{\cdot}002$).

These latter results are in line with those reported earlier by Mowbray (1953) relating to the dominance of the visual channel over the auditory with conflicting simultaneous combinations. The explanation for this visual dominance presumably resides in the nature of the auditory and visual presentation. The constantly moving visual display required an active effort on the part of the subject to the apparent

detriment of reception via the other channel. Such a discrepancy would probably not exist were it possible to accomplish simultaneous perceptions. All of the results from this experiment suggest that complex perceptions involving language cannot be effected by different sensory modes at the same time. There is no indication from this experiment what the exact limits of complexity may be. It is conceivable that very brief symbolic material, such as letters and numerals or simple words, may be handled effectively by two channels at the same time when presentation is precisely synchronized. Further research is needed on this point.

Acknowledgement

This research was partially supported by Contract N5-ori-166, Task Order I, between the Office of Naval Research and The Johns Hopkins University. This is Report No. 166-I-175, Project Designation No. NR 145-089, under that contract. The actual experimental work was done at the Psychological Laboratory, University of Cambridge, Cambridge, England, and the author is grateful to Professor Sir Frederic Bartlett and his successor, Professor O. L. Zangwill, for the facilities and assistance provided there. Dr. N. H. Mackworth of Cambridge University and Dr. W. R. Garner of The Johns Hopkins University were particularly helpful with their suggestions and criticisms.

References

Hebb, D. O. (1949). *The organization of behavior.* New York.

Mowbray, G. H. (1953). Simultaneous vision and audition: The comprehension of prose passages with varying levels of difficulty. *J. exp. Psychol.*, **46**, 365–372.

MISCELLANEA

A PROJECTION TACHISTOSCOPE

BY

L. JOHN NEWSON

(From the Department of Psychology, University of Nottingham)

It is generally agreed that a well-designed tachistoscope should satisfy the following requirements (Vernon, 1937):

1. Clear presentation of the stimulus for a measurable and variable interval of time.
2. Measurable variability of stimulus intensity.
3. Neutral exposure field, free from positive and negative after-images.
4. No variation in the intensity of the exposure field when the exposure takes place.
5. No distraction from movement of mechanical parts of the apparatus.
6. Comfortable viewing, the subject sitting in a normal relaxed position and using both eyes.

Most of the designs of which the writer is aware either fail conspicuously to fulfil one or other of these requirements, or else are so complicated in design that, though in theory they should satisfy the requirements, they still have serious deficiencies in practice. For instance, it is difficult to see how, using either the Dodge tachistoscope or the Vincent-Wolters tachistoscope, it is possible to avoid some background flicker accompanying the exposure of the stimulus; and in both cases the subject is required to peer through a hole into a box, in which the intensity of light differs from that outside, so that he almost always has a feeling of strain after the first few exposures.

The basic difficulty seems to be that, if one thinks in the conventional way of dark stimulus figure on a white ground, it is necessary to make complicated provisions for maintaining an equivalent white ground during the pre- and post-exposure periods, and for making the change-over unnoticeable. This conventional approach is abandoned in the present design, and the subject views a light stimulus-figure on a relatively darker ground. The great simplification introduced by this simple step is attributable to the fact that light can be interrupted physically, whereas darkness obviously cannot. For most practical purposes, it appears to make no difference to the subject whether he is asked to view a light stimulus on a dark ground, or *vice versa*. This is, after all, what he does every time he looks at a blackboard.

General design. The apparatus to be described is in principle extremely simple. It consists of a standard 35 mm. lantern-slide projector used in conjunction with a shutter. The slides used are negatives, which let light through only in the "figure" which constitutes the stimulus. When projected, the "ground" remains constant, simply because no additional light falls on it during exposure. Writing or drawing thus appears in white against a relatively darker background.

Slides. Stimulus material is drawn with black ink on white paper, and photographed, using a 35 mm. camera. The film is developed to make negatives, which are mounted and used as slides. For line drawings it is important to use high contrast, fine grain microfilm (as used by libraries), since only with this type of film can sufficiently contrasted negatives be obtained.

Projection screen. Using a modern high-power projector, the stimulus can be shown on a neutral grey screen at a distance of about ten feet in normal artificial room illumination. By choosing the correct shade of grey, after-image effects can be almost entirely eliminated.

Shutter. For many purposes, a camera shutter placed in front of the projector lens will be found adequate. The shutter requires no lens, and should have an internal diameter of approximately 1 inch. A good quality variable-speed shutter of the Compur type should be sufficiently accurate to allow calibration in units of approximately 1/100 sec., in a range between 1/100 sec. and 1/10 sec. Since intensity may also be controlled (preferably by using grey glass filters or wedges), this range should usually be sufficient. It is also advisable to use a heat-absorbent glass filter between the shutter and the projector, to prevent the shutter from warming up.

Where more accurate control of exposure-speeds is required, some sort of rotating disc shutter may be preferable. In practice, we have found it convenient to have a metal disc with variable aperture situated between the condenser lens and the slide carrier, and running as close as possible to the face of the slide. The disc is driven by a constant speed motor at about 1 rev. per sec., and is approximately 14 in. in diameter. This allows calibration of the exposure-time in units of 1/400 sec. or even less. The disc also rotates slowly enough to allow the experimenter to open and close an ordinary camera shutter (set to B), so that the stimulus, instead of recurring, is only projected once.

In addition to satisfying the main criteria, this apparatus has several other advantages. Its construction and operation are simple. Use is made of standard equipment, which is still available for other purposes. When investigating subjects one at a time, using a photographic shutter, each subject may operate it himself, by means of a cable release. This minimises the possibility of temporary lapses in attention or blinking during the brief stimulus presentation. The stimulus can be viewed, however, by more than one person at a time, and the apparatus is therefore suitable for demonstration purposes. Eyestrain is minimal, since the apparatus works well in normal lighting conditions. The intensity of the external light source should, however, be kept constant, since the eye is extremely sensitive to changes in the relative intensity of the stimulus. The apparatus may also be adapted for the continuous exposure of stimuli at very low levels of intensity, and can thus be used in the investigation of perceptual threshold phenomena.

Reference

VERNON, M. D. (1937). *Visual Perception.* London.

THE QUARTERLY JOURNAL OF
EXPERIMENTAL PSYCHOLOGY

Vol. VI 1954 Part 3

THE IMPACT OF VISUAL AID DISPLAYS SHOWING A MANIPULATIVE TASK

BY

S. LANER

From the Department of Psychology, University of Reading

A moving film and a filmstrip, composed of line drawings, were compared in respect of their efficacy in imparting instruction about a manual task. Their respective effects were tested by making the subjects actually perform the task. The chief aims of the experiment were, firstly, to determine the manner in which material presented pictorially is utilized in the approach to and execution of the task and, secondly, to assess the special effects, if any, of motion and photographic realism in the presentation.

The approach both to the perception of the material and to the execution of the task was found to be essentially constructive. For this reason much that is usually presented is superfluous while, on the other hand, points of principle necessary to adequate construction are often insufficiently clarified or emphasized. No significant difference between the moving film and the filmstrip presentation was found. It is concluded that aspects of visual aid material, such as motion and photographic realism, which are commonly assumed to be of special cognitive importance, do not themselves play a unique rôle.

I

Introduction

The increasing use of pictorial, in place of verbal methods of teaching and instructing has raised a number of psychological problems of both theoretical and practical interest. The widely expressed preference for pictorial methods appears to arise out of a fairly generally held assumption that media of communication which are deemed to represent "reality" with a high degree of fidelity are *ipso facto* more likely to make an impact upon the perceiver and, by virtue of this greater impact, produce quicker, greater and more durable learning effects. Although rarely expressed, the psychological implication of this assumption is that the perception of highly realistic pictorial displays is in some way simpler, more immediate, than the perception of a less "realistic" display, let alone of a display consisting of highly abstract symbols, e.g., words.

Now the motion film is typical of what is meant by a highly realistic display. If the above assumptions hold good, then the properties of continuity and represented motion should give the film a high advantage, in terms of learning effects, over both more static and discontinuous display, such as filmstrips and slides, and over verbal and textual display. The advantage should be particularly high when the "reality" represented is itself a moving and continuous one; this would appear to be the case in manual operations forming a skilled performance.

In the present investigation an attempt was made to compare a commercially produced film, showing a skilled task, with a specially prepared filmstrip in terms of relative learning effects. The commentary was controlled by keeping it essentially

the same for both types of display. One or the other type of display was shown to each subject once. Then he was required to perform the task shown. The subjects were university students and National Servicemen.

As it was one of the main aims of this study to find out whether pictorial displays do in fact make a greater impact on the perceiver, at what points the impact is greater and wherein it consists, the quantitative analysis of overall learning effects was supplemented by a kind of item analysis and by an assessment of the qualitative nature of subjects' responses. An attempt was then made to arrive at a hypothesis concerning the processes which are evoked by pictorial displays.

II

Methods and Procedures

The task.

The selection of a task suitable for laboratory treatment was governed by the availability of commercially produced film material. A number of other considerations, such as that the task should be neither too difficult nor too easy for the subjects available, and that it should be interesting and meaningful and at the same time unfamiliar to most subjects narrowed the range of possible tasks down to one: the dismantling, repair and subsequent reassembly of a sash-cord window.

The three stages of this task could be considered as each consisting of a number of discrete sub-operations: 8 in the case of dismantling, and 10 and 4 respectively in the case of the repair and reassembly stages. From a pre-view analysis of the film it appeared that the second stage of the task (repair) differed from the other two in level of difficulty, the relationships between individual actions and movements and their ultimate effects not being immediately obvious.

The display material.

For the purpose of this experiment the task-performance as shown in the film was considered to be the "correct" performance.

The film (which was produced by the Topical Film Co., Ltd.) presented the task in a completely straightforward fashion. A short description of the system on which the sash-cord window works and of its component parts was followed by a practically unabridged version of a handyman going through all the motions of dismantling the left-hand part of a window, repairing the broken cord and reassembling the window. The "camera-behind-operator" technique was used throughout. The rate of development was fairly brisk. No part of the operation was repeated, nor was any operation given special emphasis. No use was made of slow motion or of any other special effects. The commentary was equally straightforward. Possible errors were never shown visually and were at most summarily referred to in the commentary. The film has been described by an experienced producer as a medium quality one.

The filmstrip, consisting of 51 frames, was prepared in the laboratory. The intention was not to make it as closely comparable as possible to the typical commercial filmstrip, but to produce a display in which the visual element would be not only greatly reduced by comparison with the film but also lacking in the attribute of represented movement. At first it was planned to select suitable "stills" from the film itself; but the display obtained by this method turned out to be so poor in definition that it was discarded and a black-and-white version prepared instead. The "stills" for this version were obtained by tracing and inking-in enlargements of the original stills, and printing them on 35-mm. stock.

The commentary for the film and the filmstrip were both recorded on a wire recorder, a sound-film projector not being available. The wording of the commentaries for both types of display were virtually identical in phrasing and content.

Subjects and procedure.

Records of performance were obtained from 75 subjects, all of whom were tested individually. Of these 12 were university students; all of them viewed the film. The remaining 63 subjects were National Servicemen (R.A.F.), and, of these, 50 viewed the film, 13 the filmstrip.

The National-Servicemen groups consisted of personnel employed on general and clerical duties; they were in the main of medium intelligence, and their general educational background was very varied, ranging from elementary school to grammar school standard; there were also considerable differences in mechanical skill and experience. The students were generally of above-average intelligence.

Each subject was told in advance that he would be asked to perform the task after viewing. Personal data were collected immediately after the viewing, together with subjects' comments on the display. A model sash-cord window was used for the actual test, the cord being broken on the side opposite to that shown in the displays. A close record was kept of each individual performance, but timing was dropped after a try-out, as it tended to make subjects nervous and also because the method of prompting the subjects whenever their performance came to a prolonged stop rendered the time records virtually useless.

Intelligence test scores (AH4—a written intelligence test containing verbal and non-verbal items) and NIIP Spatial Relations test scores were obtained for 47 Servicemen who viewed the film and for all Servicemen who viewed the filmstrip. Similar scores for the student subjects could not be obtained.

III

Analysis of Results

1. *Overall Learning Effects.*

Preliminary tests with subjects whose results are not included in this report showed that the responses on the 23 sub-operations of the task could be conveniently classified into seven categories: correct response, major error, prompt, minor error, correct response but out of serial order, omission and importation. The categories are self-explanatory, except for the distinction between major and minor errors. This distinction was drawn on the basis of the character of an error in relation to the task as a whole, major errors being those that would have arrested the performance at some later stage or which would have led to a faulty final achievement; with this type of error the experimenter's intervention was necessary and there is, therefore, no essential difference between major errors and prompts.

Conversion into percentages of totals of correct responses and of prompts and major errors yielded two "scores" for each subject; this made possible a statistical comparison of the main features of the performance in the three groups of subjects. The statistics obtained for correct responses and for major errors plus prompts (in percentages) are shown in Tables I and II respectively.

TABLE I

Group	*Number of subjects in group*	*Mean per cent. operations correct*	*Standard deviation*
Students (film)	12	77·39	15·32
R.A.F. (film)	50	66·23	17·41
R.A.F. (filmstrip) . .	13	58·75	19·82

TABLE II

Group	*Number of subjects in Group*	*Mean per cent. major errors plus prompts*	*Standard deviation*
Students (film)	12	13·38	12·83
R.A.F. (film)	50	19·61	13·37
R.A.F. (filmstrip) . .	13	27·04	11·23

The differences between the mean per cent. operations correct of the student and the two R.A.F. groups are significant at the 5 per cent. level (the values of "t" being 2·01 and 2·42). There is no significant difference between the means of the two R.A.F. groups for correct responses. As regards errors and prompts, the difference between the student-group mean and the R.A.F. (filmstrip) group mean is significant at the 2 per cent. level; the remaining differences are not significant at the conventional levels.

The absence of significant differences between the two R.A.F. groups suggests that as regards overall learning effects the two types of display used were equivalent. The superiority of the student group must, in all probability, be ascribed to factors other than display effects.

Inspection of the two tables shows that in general the mean percentages of correct responses are rather high and the percentages of errors and prompts correspondingly low. This in turn suggests that both types of display were exceedingly effective in conveying the task. The numerical indices, however, are to a certain extent misleading for the following reasons: (*a*) the conversion into percentages of scores obtained on a task consisting of only 23 sub-operations is not unobjectionable, (*b*) the high scores are almost certainly not due to the display effectiveness alone, but at least partly to related past experience operating during the task performance, (*c*) the overall response scores tend to disguise differences in the adequacy of performance at different stages of the task which might well reflect differences in the effectiveness of the display.

2. *Performance differences within the task.*

The second method of analysis has special reference to point (*c*) above. The method was to pool seaprately the correct responses and the errors and prompts made at each of the 23 sub-operations for each whole group of subjects; these totals were again expressed in percentages. Thus if, for example, all 50 subjects in the R.A.F. (film) group performed the first sub-operation correctly, this would represent a 100 per cent. correct response for the group on this sub-operation.

The results of this analysis are represented graphically in Figures 1 and 2. Figure 1 shows the percentage of correct responses made by each of the three groups on the 23 sub-operations, Figure 2 the percentage of major errors and prompts.

Two points are clearly brought out by this analysis. Firstly, the differences found between the three groups by the preceding statistical analysis are based on

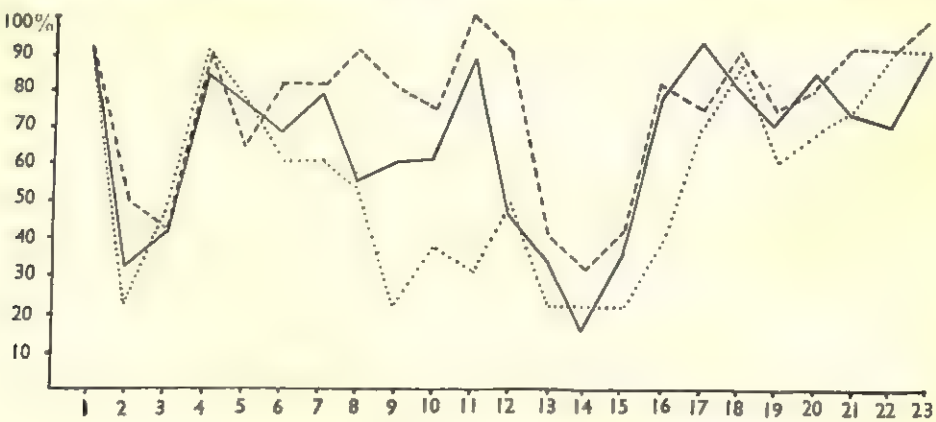


FIGURE 1

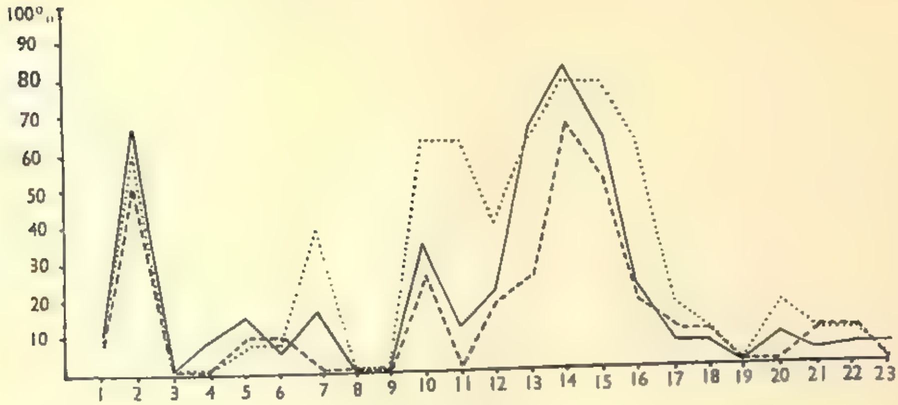


FIGURE 2

fairly consistent trends running through the whole course of the performance. The Student (film) group graph is consistently above that of the R.A.F. (film) group. This group in turn shows up consistently better than the R.A.F. (filmstrip) group. However, the advantages of any group over any other, though fairly consistent, are at most points insignificantly small.

The second point brought out by the analysis is of central importance in the present context. It is quite clear that faulty responses and errors did not occur randomly throughout the performances, but tended to pile up at certain stages of the task. This accumulation is particularly noticeable between sub-operations 9 to 18 and on sub-operation 2; it occurred with all three groups regardless of the type of display used.

3. *Relationship of performances and displays.*

The conclusions that can be drawn so far tell us about the nature of the performance rather than about the performances as related to the displays. Although it is clear that errors and prompts accumulated at certain stages irrespective of what type of display was used for instruction, it does not necessarily follow that these accumulations are due to a failure of communication by the displays; that this was so is only a plausible hypothesis. A further, qualitative, analysis might provide some further leads.

Considering first of all the qualitative nature of the errors in the initial (dismantling) and final (reassembly) stages, the majority of errors and prompts were concerned with performing an operation in a different serial order from that shown in the displays, or with omissions of sub-operations which were not of direct importance to the successful dismantling or reassembly. That alterations in serial order should have occurred is not surprising in view of the general difficulty of establishing such an order, as Bartlett (1952) noted.

Sub-operation 2, which falls into the dismantling stage, requires separate consideration. This sub-operation is that of taking the bottom sash out of the frame of the window and detaching its intact cord, thus permitting access to the second (top) sash, the cord of which had been broken. The errors that occurred here are probably due to two factors. Firstly, the sub-operation falls near the beginning of the task, and it is a common observation that subjects, when confronted with a task, are reluctant to make a move which they feel cannot be easily reversed. Secondly, the removal of the bottom sash is practically identical with the removal of the top

sash, which is also shown in the displays. In both displays no clear differentiation is made between these two nearly identical moves, which occur at different stages of the task. It is therefore likely that this lack of differentiation by the displays leads to the jumbling up of the two sub-operations in the process of perception. Clearly, this jumbling must be ascribed to a failure of adequate communication by the displays.

The quality of performances during the repair stage, where most errors and prompts occurred, was radically different from that of performances during the assembling stages. The repair stage performances frequently started with an omission of a less important operation (stretching the new cash-cord before insertion). The next two operations are concerned with the preparation of the "mouse" (a piece of lead hammered round a string which later acts as a "shoelace" for the new cord). The film groups seemed to find little difficulty here, while the filmstrip-instructed subjects fared decidedly worse. It is possible and indeed likely that this was due to the lack of continuity in action of the filmstrip.

The next sub-operation, involving the insertion of the "mouse" into the aperture above the pulley, seems to mark the point where effective communication by both displays broke down. There are several clear indications of this. For one thing, prompts were very much more frequent than errors here—in other words, subjects were completely in the dark what to do next. Secondly, subjects reported "I didn't see clearly" rather than "I have forgotten," or "I didn't remember." All in all, failure of communication, as judged by performance, occurred with 90 per cent. of all subjects regardless of the type of display.

4. *Performance and "audience variables."*

Unless a very naïve view is taken of the nature of perception, it is quite obvious that the failure or success of communication is a function of variables within the responding subject as much as of variables within the display. There is always a possibility of improving communication by altering the variables on one side or the other or on both. Since the present research is chiefly directed towards the analysis of variables in the display, our main concern was with keeping the audience variables controlled. However, the institution of preliminary controls is bound to be a somewhat futile procedure unless some foreknowledge is available regarding the significant variables. It was therefore merely attempted (*a*) to check on the homogeneity of the population with respect to "intelligence" and "spatial ability" at the close of the experiment, and (*b*) to guess at and check on any other variable that might require controlling in future experiments.

Statistical treatment of the scores on test AH4 and on the NIIP Form Relations test, which were available for 47 out of the 50 film-instructed R.A.F. subjects and for all filmstrip-instructed R.A.F. subjects, showed that between these two groups there were no significant differences on either of the tests employed, nor were any such differences found when the verbal and non-verbal parts of test AH4 were evaluated separately. This can be taken as indicative of the fact that the two samples were drawn from the same population.

To check on the possible effects of these "audience variables" the data of the 47 film-instructed R.A.F. subjects were further analysed as follows. First, the mean "correct response" percentages of subjects who scored *above* and *below* average on AH4 were compared. The same procedure was used to compare those who scored above and below average on the NIIP test. Table III shows the differences to be low and not significant statistically.

TABLE III

	N	*Mean per cent. "correct responses"*	*Difference (in per cent.)*	t
Above gp. average scorers on AH4	24	70·16	7·16	1·36
Below gp. average scorers on AH4	23	63·00		
Above gp. average scorers on NIIP	22	72·00	10·05	2·01
Below gp. average scorers on NIIP	25	61·95		

The small differences found are in good agreement with the low positive/biserial/ correlation coefficients found in a U.S. film project (Zuckermann, 1949) between scores on a film-mediated knot-tying task and the U.S. Navy General Classification (intelligence) test ($r = 0·14$) and between this task and the U.S. Navy Mechanical Aptitude test ($r = 0·30$).

Finally, there seemed to be good reasons for assuming that *related* past experience, in this instance mechanical training, experience or private interest, would affect the amount as well as the quality of performance on the task. Since during the experiment every subject in the film-instructed R.A.F. group had been asked whether he had had any mechanical training or experience or whether he was interested in matters mechanical, a statistical check, admittedly rather rough and ready, was possible. This was done by splitting the "correct response" percentages into two sets ("experienced"—"inexperienced") and by comparing the mean values. These values are shown in Table IV.

TABLE IV

	N	*Mean per cent. "correct responses"*	*Difference (in per cent.)*	t
"Experienced"	22	73·19	13·23	2·24*
"Inexperienced"	17	59·96		

* The value of t found is significant at the 0·05 level.

The significance of the difference is underlined by the fact that out of the 8 subjects in the total group who satisfied the experimenter that they had mastered the principle of the repair stage, 7 fell into the "experienced" group, and only 1 into the "inexperienced."

Although the data here presented indicate that while related past experience is a significant variable and must therefore be controlled, it should not be hastily concluded that "intelligence" and "spatial ability" (as measured by tests) are not significant variables. The scores of the R.A.F. subjects all fell within a rather narrow range, and it is possible and indeed likely that, e.g., very highly endowed subjects would (*a*) gain more information from the display, and (*b*) show more versatility in performance.

Commentary effects.

Within the framework of this experiment no direct provision was made for an estimation of the relative contributions made to learning by the visual and auditory channels of the display material. In comparing the effect of the film and the

experimental filmstrip the verbal commentary was merely kept constant for both types of pictorial display.

It was, however, decided to carry out a rough check on the commentary at two levels of assumed difficulty. At the lower level of difficulty three figures mentioned by the commentary were singled out, and subjects were asked (after the performance) if they could recall these figures. Since none of these figures could have been obtained from the pictorial display, it could be assumed that their recall would be based on a "pure" commentary effect. Table V shows the percentages of subjects in each group who recalled literally all three, two, one and no figures respectively.

TABLE V

	Figures recalled (in per cent.)			
	None	*One*	*Two*	*All*
R.A.F. (film) group	4	20	48	28
Student (film) group	—	8·5	50	41·5
R.A.F. (filmstrip) group ..	30·7	15·4	23	30·7
All subjects	8	17·3	44	30·7

From this table it is obvious that nearly 75 per cent. of all subjects tested recalled accurately two or all figures. At a comparatively low level of difficulty it is clear that the commentary is highly effective.

Under the term "high level of difficulty" we arbitrarily subsumed four items of technical terminology used in the commentary: the guard bead, the parting bead, the pocket piece and the "mouse." The data for the two R.A.F. groups only were available. The percentage of correct recall of each item are shown in Table VI.

TABLE VI

	Frequency of recall (in per cent.)			
	Parting bead	*Guard bead*	*Pocket piece*	*"Mouse"*
R.A.F. (film) group	10	10	4	44
R.A.F. (filmstrip) group ..	15	—	—	69

Only one single subject in the R.A.F. (film) group recalled accurately all the items, and two further subjects recalled 3 items correctly. In the filmstrip group no subject recalled all or even 3 items accurately.

However, there was no dearth of "substitute" technical terms, and practically no subject was prepared to admit that he recalled no names of technical parts: 8 different substitute versions were volunteered for the guard bead, 23 different substitutes for the parting bead, 8 for the pocket-piece. The most interesting and the most amusing transmutations occurred with the term "mouse," which was recalled accurately most often. The substitutes offered included (apart from descriptive suggestions) such terms as "rat," "cat," "mask" and even "fish."

Most subjects pointed out that in viewing the display and trying to get an idea of what is going on, there is no time to "register" names of parts as well as actions. In fact, some subjects objected that the technical terminology tends to confuse,

e.g., "The names baffled me and put me off." The conclusion that little emphasis should be put on terminology where the chief aim of the display is to convey a coherent account of some action or actions is supported by similar conclusions reached in U.S. research to the effect that technical terminology "does not appear to facilitate the learning of an assembly skill and may actually interfere with such learning" (Jaspen, 1950).

The "impact" of pictorial material.

The most obtrusive qualitative feature of the subjects' performances in these experiments was their manifestly constructive nature. There was an almost infinite variability and individuality not only in the actions and movements, but in the very approach adopted by any given subject, so much so that to an outside observer it might occasionally have appeared doubtful whether the performances were based on the same display or indeed on any display at all. With practically all subjects the performance was more clearly related to the actual task than to the display seen. Not by any stretch of imagination could the performances in any of their aspects be characterized simply as reproductions of the movements and actions shown on the screen. This was particularly striking at the more difficult stages of the task, where subjects often suspended action altogether, apparently to improvise some more or less satisfactory response, or where a wrong response or series of wrong responses tended to occur.

The observed constructiveness of performances as a whole and of its component movements and movement patterns suggested that very little reliance was put by subjects in type of recall (whether expressed in words or actions) based on some sort of serial photographic imagery. Verbal reports, too, offered no evidence that —at any rate with adult subjects—such imagery is either expected to occur, relied on, or utilized. May's hypothetical subject (1946) who carried away with him a "mental film" which he later "reviews" and imitates is less hypothetical than imaginary. This is not to say that imagery does not occur; but it seems that what imagery occurs is itself constructive in nature.

These observations in turn suggested that the "impact" of pictorial material during perception is not by any means simple and immediate, if by simple and immediate is meant that faithful copies of pictures are in some way transferred from the display into minds or brains; and it is precisely this idea which underlies many of the claims made for visual-aid instruction. Although it may still be true that the perception of a highly realistic display differs in some respects from the perception of more symbolic displays, the present experiment contained no evidence for the existence of some basic divergence from the general principles held to govern perceptual processes (Vernon, 1952). On the contrary, there was every indication that just as the performances could be described as series of individually constructed responses, so too can the perceptual process during viewing. The adequacy of performance constructs, moreover, appeared to be intimately related to the success or failure of the perceptual constructions, on which the former were almost certainly based.

The constructive perceptual responses seemed to consist in selecting from the display (not necessarily from the pictorial element of the display only) a number of cues to action (i.e., dominant details) and in organizing these within a general framework or setting which itself results from an integration of actually presented data with a schema of past experiences, brought into the viewing situation. For such an integration to occur it is important both, that the data should be adequately presented—adequacy of presentation and fidelity of representation not being

synonymous terms—and that the subject should have a fund of appropriate related experiences to draw on. The present experiment provided evidence to show what happens when inadequacies occur either in presentation or in past experiences.

Many subjects complained that the data were inadequately presented, particularly during the more complex stages of the task. These complaints were frequently framed in such terms as: "The general principle is not made clear enough" or "The construction of the window was not clear." Evidently the majority of subjects were not able to relate the series of actions shown in the displays to each other at the more complex stages, where the display gave no indication how they fitted into the general "system" or how they contributed to the overall aim of the task. It seems that only where the series of actions can be referred to the overall aim can the subject face the task with a more or less clear plan of action. As will have been evident from the quantitative analysis, 90 per cent. of all subjects approached the task without a plan sufficiently articulated to permit satisfactory performance at more difficult stages.

As regards adequacy of related past experiences, it has been shown in the quantitative analysis that such past experiences are probably significant variables. This is further borne out by the case of a subject almost entirely lacking in such perceptual-motor schemata. He could not carry out successfully more than two sub-operations of the task. In the subsequent report he stated: "Presuming it's somebody like me, you would need previous knowledge of tools and such things. I have never used tools before. But I remembered the lead at the end of the string, it was rather like a shoe lace. I had had that experience before and in a way anticipated what he was going to do in the film." Significantly, too, he maintained that "looking at the film is too passive," which probably signals the failure of keeping up perceptual responses.

In view of the preceding argument it further seems likely that the 10 per cent. of subjects who performed correctly even at the more complex stages of the task had an inordinately extensive repertoire of related past experiences to draw on, which helped them to provide for themselves a general setting despite inadequate presentation of data in the display.

That the subjects first and foremost aim during the first showing of a display is to construct the data given into a general setting is also borne out by many reports, which are mostly phrased as subjective estimates of the adequacy of achieved settings: "I have gathered the general idea"—"Easy to pick up the general idea"—"I have a rough idea of the main things." The setting, which appeared to form the basis of performance constructions, are apparently not always established consciously, nor do they necessarily operate consciously. This is reflected in such remarks as: "It comes to you as you are doing it." "It sorts itself out as you go."

As regards what we have termed "cues to action," these need not be related, and in most cases do not appear to be related, to the movements and movement patterns portrayed in the display. In fact, subjects reports as well as observations of their performances suggested that little heed is paid to movements and action details, at least where the display is shown for the first time and is not repeated. What seemed more important was the significance, the meaning of movements and actions in relation to the manipulated objects—the effects they achieved. It further seemed that both the number of cues selected and their fitness for use was related to the adequacy of achieved settings.

The nature of the perceptual process during visual-aid viewing is also partly revealed in the rareness with which the serial order of operations shown in the display was followed in actual performances. Not only was there never any suggestion that

the subjects were trying to copy deliberately the display order, but the general behaviour of subjects indicated that they felt they could not be expected to attempt this after a single showing.

Finally, it is necessary to qualify the references to the "display" made in the above context. Since the instructional media used in this experiment were in both instances combinations of pictorial material and verbal commentaries, it is obviously impossible to refer the observed effects to the pictorial element only. Although it is impossible to claim this with any certainty on the bases of the present experiments, it seemed that in so far as the instruction conveyed involved relationships, the comprehension of these relationships was not directly imparted by the pictorial element, but rather by the spoken commentary. It seemed that for such comprehension to occur, the relationship must be formulated verbally either by the subject himself or by the commentary. There is little evidence that subjects can do this. This point, which is of great importance, requires further investigation.

Conclusions.

The conclusions that can be drawn from this pilot investigation must not be regarded as anything but tentative at this stage and this applies with even greater force to any generalizations which suggested themselves as a result of the experiments described. Furthermore, it cannot be assumed that the findings can be extended from the tasks studied to all other tasks or to the communication of information not intended for giving instruction in a task-performance. Within its own limitation the present study suggests these points:

The quantative analysis of the performances mediated by the two types of display (film, filmstrip) did not indicate that continuity in pictorial representation or the inclusion of movement representation (film) necessarily produces better result in terms of correct responses. There were no significant differences between the overall scores of the groups of subjects that could be attributed to specific display effects. The analysis of responses at each sub-operation of the task showed that only at one point could increased learning effects be attributed to continuity in the display, while there was no evidence for any advantage due to motion representation; it is, however, possible that any advantages that might have occurred as a result of motion representation were cancelled out by the restricted temporal availability of the film passages at any given point. It is significant that the performances observed were least adequate at those stages of the task, where the actions shown in the display were related to each other in a more complex manner.

This last point, when considered together with the qualitative character of subjects' performances, virtually excludes the notion that subjects based their performances on photographic images or series of such images. Although subjects frequently testified to the occurrence of imagery, as might have been expected, these images were usually fragmentary and certainly bore no point to point relation to the displays. The performances themselves gave the impression of being constructions devised to meet the requirements of the task.

From these observations it was argued that the perceptual processes that occur during display-viewing are also essentially constructive responses which result in the integration of display data with the subjects own related perceptual-motor schemata into settings, and in the extraction of certain dominant details in the form of "cues to action." With reference to the pictorial displays, there was nothing to suggest that these are either stored as memory-images or that they could furnish ready-made meanings. Such meanings are not supplied by the display, but "read

into the display" by the perceiver, though the display might in some instances narrow down the range of possible meanings derived.

From the practical point of view a faithful step-by-step and movement-by-movement demonstration of a task by pictures seem extraordinarily unimportant, particularly if it is achieved at the expense of an adequate explanation of how various actions are related to each other, to the object manipulated and to the final aim of the performance. It appears that the subject often can and usually does construct for himself a great number of necessary actions, provided he has grasped certain crucial features, while adding nothing substantial. Moreover, where these crucial features involve relationships, it is likely that the verbal formulation on the commentary rather than the pictorial element contributes to adequate presentation.

All-in-all, fidelity of representation emerges as a misleading notion in the construction of visual aids to present a manual task. It is adequacy of presentation that should be aimed at and adequacy of presentation is not synonymous with fidelity of representation. What constitutes adequacy of presentation in visual aids is a matter for further exploration.

My grateful acknowledgements are due to the Trustees of the Nuffield Foundation for the financial support given this research and to the University of Reading for providing laboratory facilities. I am indebted to Professor R. C. Oldfield, who has sponsored and steered this research from the start and to Dr. M. D. Vernon for help and advice at all stages. I also wish to thank the Director-General of Medical Services, Air Ministry, the Air Officer Commanding F.T.C. and the Officer Commanding H.Q. (Unit) R.T.C., R.A.F., for this kindness in arranging to make subjects available. Of the many others, who through their co-operation are helping this research along special mention must go to Mr. John Shearman of the Film Section, Iraq Petroleum Co., Ltd., and to Miss Mary J. Locke, Technical Information Specialist at the Office of the U.S. Assistant Naval Attaché for Research. Last but not least my thanks are due to Mr. C. W. Murfin for his technical assistance and especially for constructing the dummy sash-window.

References

Bartlett, F. C. (1932). *Remembering*. Cambridge.

Jaspen, N. (1950). Technical Report SDC 269-7-17, Pennsylvania State College Instructional Film Research Program.

May, M. A. (1946). The psychology of learning from demonstration films. *J. educ. Psychol.*, **37**, 1-12.

Vernon, M. D. (1952). *A Further Study of Visual Perception*. Cambridge.

Zuckermann, J. V. (1949). Technical Report SDC 269-7-4, Pennsylvania State College Instructional Film Research Program.

AN ILLUSION OF MOVEMENT COMPLEMENTARY TO THE HORIZONTAL—VERTICAL ILLUSION

BY

R. L. REID

(From Canterbury University College, Christchurch, New Zealand)

Blindfolded subjects moved a stylus held in the hand over a standard distance of 4·5 ins. in a given direction. They then attempted to move the same distance in a direction at right angles to the first. Eight combinations of movements were investigated. The results reveal an illusion such that the extent of movements to left or right across the body is underestimated, while the extent of movements towards or away from the body in the mid-line is overestimated. The illusion applies to speed as well as extent of movement. Movement up or down in a vertical plane is equivalent to movement towards or away from the body in a horizontal plane.

The interaction of this illusion with the well-known horizontal–vertical illusion of visual perception explains a failure to find any net illusory effect where lines visually displayed in different orientations were matched for length by unseen movements in similar orientations.

Whether the visual and movement illusions simply co-exist or whether they are functionally related is not yet clear.

I

Introduction

It has long been known that observers will match a line of given length displayed vertically with one of greater length displayed horizontally. Effects related to this horizontal–vertical illusion also occur. For example, vertical movement appears faster than objectively equal horizontal movement. Since vertical distances are visually exaggerated this is to be expected.

At first glance it seems curious that human beings should permanently be subject to a distortion of visual extents according to their orientation. Distortion could only impede adaptive functioning. However, the type of environment that might be supposed to have shaped human evolution would not very frequently present the organism with situations like those that arise in the laboratory when the horizontal–vertical illusion is being measured. The procedure requires subjects to compare and make judgements about the relative lengths of two lines. This is highly artificial. The primary function of visual perception must be regarded as determining movement in relation to objects seen. With this point of view in mind an experiment was designed in order to find out what happens when subjects estimate the lengths of lines in different orientations by making matching movements rather than judgments.

The subjects sat 6 ft. away from a matt black surface on which was placed a strip of white celluloid $3 \times \frac{1}{8}$ in. that could be set vertically or horizontally. They provided estimates of its length by drawing lines in corresponding orientations on a pad of paper. A screen prevented them from seeing either their hand movements or reproductions. The line was presented four times in each orientation and on each presentation the subject produced five estimates. To maintain a fresh approach to the problem subjects were led to believe that on each presentation the line would vary in length. They were told that the variation would be so small that they would not detect it visually and should not try to do so, but that their drawings might provide evidence that they could "register" the differences in length correctly.

The results failed to show any systematic effect of orientation of the line. For a group of 15 subjects the mean length of reproductions with line horizontal was

2·40 in., and with line vertical 2·42 in. The difference of 0·02 in. amounts to less than one-third of its standard error ($t = 0·28$).

Similar results were obtained whether subjects made their drawings on a horizontal or on a vertical pad of paper, except that variability was increased when the pad was vertical. With the pad horizontal, the vertical line was drawn by making a movement towards or away from the body along the mid-line. Subjects stated that this movement "felt the same as" vertical movement up or down.

Provisional acceptance of these results at their face value led to a direct test for an illusion of hand movement that would counteract the visual horizontal–vertical illusion.

II

Experimental Arrangements

Subjects

Twelve students, eight male and four female, took part in the experiment. All subjects were right handed.

Apparatus

A 12-in. square of thin glass was set in a frame with edges of smooth hard wood. Graph paper was laid beneath the glass so that distances outwards from each corner could be read in inches and tenths. A thin metal stylus was used by subjects in tracing along the edges of the frame.

Procedure

In all trials the subject worked blindfolded sitting at a table with the board lying flat in front of him. The stylus, held in the right hand, was guided to a position 4·5 in. from a corner of the board. The subject traced along the edge from this position until the stylus reached the corner. This established the standard distance. He then moved along the edge at right angles to the original direction and stopped when he judged he had moved the same distance.

Eight combinations of movement were used. They are shown in Figure 1 with arrows indicating direction. Subjects knew in advance which combination was required. The order of performance was varied at random from subject to subject with the restriction that A and B combinations should alternate. Ten trials were carried out by each subject for each of the eight combinations.

FIGURE 1

Combinations of movement. In all cases movement in the first direction was limited to 4·5 inches.

Subjects were told that the accuracy of hand movement in different directions was being assessed. The procedure was then demonstrated and followed by a number of practice trials that varied according to individual requirements. During both the practice period and the experiment proper, subjects were allowed to see the board only when a new movement was being demonstrated. No guidance was given on methods that might help subjects to make their estimates, but they were told that if they did adopt some method during the practice period they should try not to change during the course of the experiment.

III

Results

Table I shows mean lengths of estimating movements and constant errors for each of the eight combinations. An analysis of variance into three components is given in Table II. The hypothesis that the eight means are drawn from the same population can be confidently rejected. The F-ratio of 15·56 exceeds the value of 5·78 required for significance at the 0·01 level. The difference between pairs of means required for 0·01 significance using the t-test is 0·72. Each of the four means for A combinations of movement differs significantly from each of the four means for B combinations. In addition, A4 differs significantly from A1 and A3.

TABLE I

A			*B*		
Movement combination	*Length of estimate*	*Constant error*	*Movement combination*	*Length of estimate*	*Constant error*
1	5·04	+0·54	1	4·24	−0·26
2	5·47	+0·97	2	4·03	−0·47
3	5·24	+0·74	3	4·16	−0·34
4	6·00	+1·50	4	4·20	−0·30
Mean A	5·44		Mean B	4·16	

Standard Movement .. .. 4·50
Geometric Mean of A and B .. 4·75
All entries in inches

The means of the four A and the four B combinations are given in Table I. The geometric mean of these is 4·75 in. showing a general overestimation of the standard amounting to 0·25 in. The geometric mean was calculated since the illusion can be assumed to operate proportionally. Since the standard necessarily came first on all trials this can be regarded as a time error.

TABLE II

Source	*d.f.*	*Sums of squares*	*Variance*
Conditions ..	7	4,570	652·86
Subjects ..	11	2,999	272·64
Residual ..	77	3,170	41·17
Total	95	10,739	

F = 15·86 (7 and 77 d.f.)

Computation from mean estimates of each subject for each combination of movements.

IV

Discussion

No plausible hypothesis can be suggested to account for the significant difference between the mean of A4 and the means of A1 and A3. Otherwise the general pattern of results shows that movements are phenomenally equal when those made

to left or right are objectively greater than those made towards or away from the body. An illusion of movement clearly occurs under the conditions as described.

After the end of the experiment subjects were asked to describe the method they had used, if any. As expected, most reports were vague, but seven subjects said that they had tried to keep speed of movement constant while estimating time of moving. Three of these at least carried on continuous subvocal counting. Their results were no different from the results of others who did not mention speed in their reports. Hence there are grounds for believing that apparently equal speeds of hand movement are in fact faster when the movement is to left or right than when it is towards or away from the body. This effect may be similar to one that occurs in vision where horizontal and vertical movement are seen as equal when the horizontal movement is objectively faster. The tendency to increase speed on movement to left or right may have some practical importance.

It should be noted that all movements took place in the horizontal plane. No fundamental difference occurs when measurements are made with the board set vertically in front of the subject. The illusion still holds with vertical up or down movement substituted for outwards or inwards movement. In the experiment described, the horizontal position was used because it provides data of greater reliability than the vertical position. Vertical movement appears to be equivalent phenomenally and as far as illusory effects are concerned to towards and away movement. This is not surprising. In drawing, people seem to find no great difficulty in using a horizontal pad of paper on which they register vertical extents by towards and away movements. On the other hand, making a drawing of a scene representing verticals by lines in any other orientation is awkward. A two-dimensional representation of three-dimensional space involves the projection of depth as vertical separation.

The visual horizontal–vertical illusion may show the same effect in the horizontal as in the vertical plane, but this would be very difficult to check experimentally. With both lines in the horizontal plane and the line of sight horizontal, foreshortening of one of the lines would occur but would be partially masked by constancy effects. The net result would be difficult to predict and the experiment would be trying to subjects who would have to make judgments based on peripheral vision.

The movement illusion results in a figure like that of the horizontal–vertical visual illusion. Knowing the facts of both illusions we can explain the negative result of the experiment described earlier in which visual stimuli were matched by movements. It seems reasonable to postulate that the tendency for a vertical extent visually perceived to be judged greater than a horizontal is counteracted by a tendency for vertical and vertically related movement to appear greater in kinesthetic perception than horizontal movement. To try to elucidate the process, we can think of it like this. If a person looks at a square he sees a rectangle of slightly greater height than breadth (visual illusion). If he attempts to draw such a rectangle blindfold he will produce a square (movement illusion). Thus two illusions cancel each other's effects. Of course, this is a way of explaining the data rather than an attempt to describe any processes that actually occur.

The movement illusion as shown above is greater than the horizontal–vertical illusion would be with an equivalent group of subjects, but against this inequality may be set the fact that through everyday experience subjects may have a little practice in equating lines as seen but are likely to be completely naïve to the task of equating movements.

The visual illusion is loosely named. It has never been shown to have any connection with the horizontal or the vertical in a strict sense. The illusion was

measured by the author (*a*) conventionally with subjects sitting upright and (*b*) with the head or the whole body in a horizontal position. A few subjects also made judgments lying on their backs and looking up or lying on their stomachs and looking down. The results showed that the illusion was determined by the orientation of head and eyes to the test figure. There was no indication that the gravitational field or the known horizontal and vertical lines prominent in the room in which the experiment was carried out had acted in any way as determinants of the illusion. The illusory effect was seen to depend entirely on the relationship between eyes and test figure. Provided that the eyes are in their normal relationship with the rest of the body, both visual and movement illusions occur with reference to the same system of co-ordinates.

The ability to make quick accurate movement in response to objects seen has a definite survival value to any organism. Although ordinarily such movement is under visual control yet feed-back processes take time and are best reduced to the minimum. It is tempting to suppose that the two illusions are not merely co-existent but that one is primary and the other a compensatory mechanism that may be established through learning due to its utility. Recently psychologists such as D. O. Hebb have stressed the extent to which learning may enter into apparently simple perceptual processes. Studies of children show that gross errors in reaching out for seen objects occur up to five or six months and suggest that the ability to move accurately in a visual world is established by a long-term process of learning.

Recent work on the visual illusion has not made its basis any clearer. On the one hand, Motokawa (1950) using the electrophosphene effect has provided some evidence that it can be attributed to retinal functioning. On the other hand, a measurement of the illusion by Pollock and Chapanis (1952) at each of the 10-degree positions between 0° (horizontal) through 90° (vertical) to 170° shows that lines tilted to the left of vertical look longer than those tilted to the right. Since the structure of the two retinae is symmetrical and not repetitive these results must be regarded as conflicting with Motokawa's, unless we can believe that judgments of length are based only on the stimulation of one eye.

Experiments in progress are designed to find out whether the visual illusion is derived from the movement illusion. Put in its crudest form, the hypothesis would be that a vertical line looks longer than a horizontal because movement in relation to it would feel longer than horizontal movement. The question is of interest for any light it may throw on the perceptual integration of different sense modalities.

References

Motokawa, Koiti (1950). Field of retinal induction and optical illusion. *J. Neurophysiol.*, **13**, 413–426.

Pollock, W. T. and Chapanis, A. (1952). The apparent length of a line as a function of its inclination. *Quart. J. exp. Psychol.*, **4**, 170–178.

A FURTHER STUDY OF THE PENDULUM PHENOMENON

BY

K. R. L. HALL and A. E. EARLE*

From the Departments of Psychology, Barrow Hospital and University of Bristol.

A first series of experiments had demonstrated certain conditions eliciting or inhibiting a "pendulum" phenomenon in the visual perception of apparent movement. The present study consists of five further variations designed to show more clearly conditions of occurrence and non-occurrence of this type of movement. The main findings are:

(1) Altering the axis of display to vertical significantly reduces the frequency of pendular-movement perception;

(2) Altering the position of metronome from behind to the side of the visual display, gives results almost identical with those where the metronome was inaudible, but, when the metronome is illuminated in this position, all forms of movement perception are reduced, and no pendular movement is reported.

The results for all the ten conditions, including the five of the first series are summarized, and the following possible factors are discussed: past experience, physiological nystagmus, and intervening adaptation. All three may be required to account for the perceptual phenomena under investigation and the dichotomizing of explanations into "experiential," or "physiological," appears to be arbitrary and inconsistent with the complexity of the observed facts.

I

Introduction

In a previous paper (Hall, Earle and Crookes, 1952), certain conditions for the occurrence or non-occurrence of an arcuate type of apparent movement between two alternating lights were described. With a metronome directly behind the display screen and producing alternating light signals synchronously with the auditory rhythm, the impression of movement of a light along a curved path was very strong. Experimental analysis of some of the conditions facilitating or inhibiting the phenomenon showed it to be complexly determined by, amongst other factors, the circular shape of the lights, their rhythmical presentation, and the warming and cooling of the lamp filaments.

In these experiments it had been noticed that, in addition to affecting the path which the movement seemed to take, certain conditions of stimulus presentation seemed to facilitate the perception of any kind of movement. This general effect was, however, not analysed further because it was only secondary to the main effect on the path of movement, but it appears that Werner and Zietz (1928) had shown intersensory effects which are somewhat similar to this. For example, they used a rate of succession of the visual stimuli which was below the accepted optimum for movement perception, and accompanied this by an auditory rhythm composed of clicks at a rate slightly faster than that of the visual stimuli. Some of their subjects reported definite movements in the visual field. Using an irregular series of taps, however, most subjects experienced a "disintegration" of the visual movement.

There are, then, two closely related problems requiring further experimental investigation. First, the main hypothesis arising from our previous experiments was that the path of seen movement must be primarily determined by a complex interaction of visual display conditions, and secondarily by reference to models derived from past experience. Our results were not conclusive, and it remains for us

* In receipt of a grant from the Medical Research Council.

to carry out further variations which may clarify step by step the nature of the original complex pendular phenomenon. Secondly, although we had noticed in our experiments that the total frequency of movement responses of all categories tended to vary in certain conditions, we did not examine this statistically. We shall, therefore, make some appraisal of this "facilitatory" effect in relation to both our present experiments and those we have already reported.

II

Method and Procedure

The subjects were, as previously, doctors, nurses, university students, administrative and laboratory staff, of both sexes, the age range being 20–35, with two subjects above the upper limit. Different subjects were used for each experimental condition, except when transfer effects were being specifically investigated, and only two had any previous knowledge of gestalt perceptual experiments.

The apparatus was identical with that used in the previous experiment, the display unit consisting of two flash-lamp bulbs set on and off alternately by a metronome, the time intervals and exposure times being the same as before to give a total event: Light 1 + Interval + Light 2 = 550 m/s.

For all the conditions of the experiment, the method of presentation and instruction to the subject was identical with that of the first experiments, the observations always being carried out in a darkened room. The subject was required simply to report his impressions, and no indication was given as to what he might expect to see.

III

The Experiments

In our previous investigation, the results of five experimental conditions were reported, and the present series of seven further variations are, therefore, numbered to show continuity and avoid confusion in back reference.

Condition 6.

The display box was set up on end, so that the two lights appeared vertically one above the other, instead of on a horizontal axis as was the case in every other experimental condition. In all other respects—viewing conditions, stimulus-shape, distance between stimuli, timing values, and the audible presence of the metronome—stimulus conditions were identical with those of Condition 1. Ten subjects were used: their reports are summarized in Table V.

TABLE V

Frequency of Four Categories of Movement Perception for Ten Subjects under Condition 6

Distance apart in cm.	*Vertical curve*	*Horizontal curve*	*Straight*	*Two-light movement*	*Total movement*	*Total no-movement*
18·5	1	0	2	0	3	7
9·5	1	1	6	0	8	2
6·5	1	3	5	0	9	1
Total	3	4	13	0	20	10

In this Table, as in all the others, the category "Vertical Curve" refers to curved apparent movement between the lights and in the same plane as the surface of the display-screen. "Horizontal Curve" refers to curved apparent movement between the lights and in a plane at right angles to that of the screen. In other words, these categories maintain a constant relationship to the main axis of the display, even though, in the one case of Condition 6, that axis has been changed from horizontal to vertical.

It will be seen that, in comparison with Condition 1, the perception of pendular movement has been drastically reduced, while that of straight movement has been correspondingly increased (comparing frequencies of straight + horizontal curve with vertical movement for the two conditions, $\chi^2 = 22{\cdot}9$, $P = < 0{\cdot}001$). Total movement has also been reduced ($\chi^2 = 6{\cdot}66$, $P = < 0{\cdot}01$).

The shift in movement perception from predominantly pendular in Condition 1 to predominantly straight in Condition 6 may be due to any of three possible factors:

(i) Differences in the organization of complex stimuli in the visual primary projection system according to whether the display axis is horizontal or vertical.

(ii) Differences in facility and track of eye-movements or physiological nystagmus in the horizontal or vertical direction.

(iii) Differences in the likelihood of the hypothetical clock-pendulum model being evoked, explicitly or otherwise, when the display axis is changed to vertical. In this respect, it may be noted that the three "pendulum" reports under this condition were all given by one subject, who spontaneously stated that she was trying to think how the display was being produced, and quickly decided that the underlying mechanism must be a metronome fixed so as to work on its side.

Condition 1 A.

Seven of the subjects of Condition 6 went on to observe the phenomenon under the original Condition 1 (horizontal axis) at the shortest distance between lights. Of these seven, four continued to report straight movement, two reported curved movement on a horizontal plane, and one now reported pendular movement. Thus the general effects of Condition 6 appeared in these subjects to have transferred to Condition 1. This finding is in accord with the effects of experimentally-induced set upon the phenomenon of apparent movement noted by Wertheimer (1912) (quoted by Vernon, 1952, p. 178).

Condition 7.

Comparison of the effects of Condition 1 and Condition 2 indicated that the *audible* presence of the metronome, as distinct from the visual rhythm with which it actuated the lights, tended to increase the frequency and degree (depth of curve) of perceived arcuate movement. In Condition 7, we set out to see what would be the effect of making the metronome not only audible, but also visible to the subject during the period of observation. The metronome was placed clear of the display unit, about three inches to the right of it from the subject's point of view. A screened flashlamp bulb illuminated the metronome with a dim reddish light which was not sufficiently bright to show up any other feature. Electrical connections were such that, as the metronome arm inclined to the left, the left-hand light, and when to the right, the right-hand light came on. Otherwise conditions were identical with Condition 1. The reports of the five subjects who observed under this condition are summarized in Table VI.

TABLE VI

FREQUENCY OF FOUR CATEGORIES OF MOVEMENT PERCEPTION FOR FIVE SUBJECTS UNDER CONDITION 7

Distance apart in cm.	*Vertical curve*	*Horizontal curve*	*Straight*	*Two-light movement*	*Total movement*	*Total no-movement*
18·5	0	0	1	0	1	4
9·5	0	0	2	0	2	3
6·5	0	1	3	0	4	1
Total	0	1	6	0	7	8

The visible presence of the metronome, it seems, leads to a marked reduction in the probability of any kind of movement being seen, and eliminates the pendulum effect altogether. It is possible that this may be a result of the illuminated metronome acting as a distracting element, and causing the observers to keep shifting their region of fixation between it and the screen. As in the other conditions, however, they were told simply to observe what was taking place on the screen, no reference at all being made to the metronome. At any rate, the visible presence of the metronome did not serve to reinforce any hypothetical models of pendular systems which may have been operative in the subject's perceptual processes at the time.

Condition 8.

We noticed that from the observer's position the sound made by the metronome was louder under Condition 7 owing to the fact that the metronome was not screened by the display box. It seemed that this accentuated auditory component might have contributed to the inhibitory effects noted under Condition 7. As a control for this, six subjects observed the display with the metronome in the same position as for Condition 7, but without the metronome being spotlit. When the room light was on the metronome was screened from the subject's view. Thus, apart from the position and loudness of the metronome, the display was identical with that of Condition 1. Table VII shows the distribution of the subjects' reports.

TABLE VII

FREQUENCY OF FOUR CATEGORIES OF MOVEMENT PERCEPTION FOR SIX SUBJECTS UNDER CONDITION 8

Distance apart in cm.	*Vertical curve*	*Horizontal curve*	*Straight*	*Two-light movement*	*Total movement*	*Total no-movement*
18·5	2	0	1	0	3	3
9·5	3	1	1	0	5	1
6·5	2	0	2	0	4	2
Total	7	1	4	0	12	6

Pendular movement reappears, but with only half the relative frequency observed under the original Condition 1 (the difference is not significant at the 0·05 probability level, χ^2 being 2·03, correction for continuity being applied). Total movement is also significantly reduced ($\chi^2 = 4{\cdot}0$, $P = < 0{\cdot}05$). Why the slightly increased loudness of the coincident metronome beat should exert this inhibitory effect on the phenomenon is not clear, but it is possible that the auditory background in this condition is sufficiently sharp and accentuated to modify the general figure-ground relationship between visual and auditory stimuli which prevails in other conditions.

Condition 9.

One of the most marked effects in the earlier experiments had been the predominance of straight movement in Condition 3, when the lights were square in shape and aligned horizontally on the main axis of the display screen. In Condition 9 we examined the effect of turning the two squares through 45° so that their horizontal diagonals were in line and parallel to the main axis of the display. Otherwise the conditions were identical with those of Condition 3. A résumé of the six subjects' reports appears in Table VIII.

TABLE VIII

FREQUENCY OF FOUR CATEGORIES OF MOVEMENT PERCEPTION FOR SIX SUBJECTS UNDER CONDITION 9

Distance apart in cm.	*Vertical curve*	*Horizontal curve*	*Straight*	*Two-light movement*	*Total movement*	*Total no-movement*
18·5	2	0	0	0	2	4
9·5	5	0	0	0	5	1
6·5	2	2	2	0	6	0
Total	9	2	2	0	13	5

Compared with Condition 3, the proportion of total movement remains about the same (72 per cent. for Condition 8, and 66 per cent. for Condition 3), but vertical curve rather than straight is now very much the predominant perception ($\chi^2 = 12{\cdot}42$, $P = < 0{\cdot}001$).

Condition 10.

An initial study of the effect of interrupted rhythm was made in Condition 5, when it was found that making the time interval between stimulus 2 and stimulus 1 longer than the standard time-interval between stimulus 1 and stimulus 2, tended to reduce the perception of vertical-curve in favour of straight and horizontally-curved movement. There the longer time-intervals had been 1·5 seconds and 3 seconds. Ten subjects observed at both intervals and at two distances between lights (9·5 and 6·5 cm.). In Condition 10, five subjects observed the display at the three usual inter-light distances and with the longer time interval constant at 6 seconds. As in Condition 5, the lights were actuated by a cam arrangement. The distribution of Group 10's reports appears in Table IX.

This group provides the smallest proportion of total movement responses of any in the whole series of experimental conditions. What movement is observed is straight.

TABLE IX

Frequency of Four Categories of Movement Perception for Five Subjects Under Condition 10

Distance apart in cm.	*Vertical curve*	*Horizontal curve*	*Straight*	*Two-light movement*	*Total movement*	*Total no-movement*
18·5	0	0	1	0	1	4
9·5	0	0	0	0	0	5
6·5	0	0	1	0	1	4
Total	0	0	2	0	2	13

At the bottom of Table X, in which a summary is made of observer's reports for all the main experimental variations, the data derived from Conditions 5 and 10 are arranged in such a way as to make the two groups' reports more directly comparable. Group 10's reports at the 18·5 cm. distance between lights are eliminated. The distribution of the remaining reports is tabulated in three rows, one row for each time-interval between pairs of lights (1·5, 3 and 6 seconds) irrespective of distance between lights.

Condition 11.

An attempt was made to examine the effects of fatigue as a possible factor in the phenomenon. Five subjects were observed under conditions equivalent to those of Condition 1 at the shortest inter-light distance. They continued to report for a period of fifteen minutes, speaking into a microphone which was coupled to a wire-recorder. The record was transcribed and analysed to see if the different types of movement perception appeared in any particular order, or were distributed in any particular way through the period of observation. No systematic tendency could be detected, there being neither any evidence for more rapid change from one type of movement-perception to another during the test period, nor for either straight or curved movement to appear first in sequence.

IV

Summary of Results

A summary of all subjects' reports, except those of Condition 11 (fatigue effect) and the "check" Condition 1A is presented in Table X.

In order to simplify tabulation, the factor of inter-light distance has been ignored, and the figures for each condition refer to the total incidence of particular movement-categories for all distances within that condition. The number of reports falling into each movement-category within any condition is expressed as a percentage of the total number of reports for that condition. Under the heading "Summary of Conditions," the variation from Condition 1 is in each case italicized in order to simplify reference to the Table. The movement-categories are those which have been used throughout the series of experiments.

In considering the "interrupted rhythm" conditions (5A, 5B and 10), it should be noted that the same subjects were used for both parts of Condition 5, while the subjects of Condition 10 had seen the display with the 18·5 cm. inter-light distance, though their reports on this, for reasons explained earlier, are not here included.

For Condition 6, as throughout the Table, "Vertical curve" means a curved path of movement seen in the same plane as the display screen. "Horizontal curve" means a curved path of movement in a plane at right angles to that of the display screen.

TABLE X

SUMMARY OF RESULTS.

FREQUENCY OF FOUR CATEGORIES OF MOVEMENT PERCEPTION, EXPRESSED AS PERCENTAGE OF TOTAL

Summary of conditions	*No. of Ss*	*Vert. curve*	*Horiz. curve*	*Straight*	*Two-light*	*Total movement*	*Total no-movement*
1. Circular, audible, met. not visible behind display, 3 dists. apart.	10	77	0	13	3	93	7
3. *Square*, rest as for 1.	10	3	0	64	10	77	23
9. *Diamond*, rest as for 1.	6	50	11	11	0	72	28
2. *Inaudible*, rest as for 1.	10	40	7	20	3	70	30
7. *Met. visible at side of display*, rest as for 1.	5	0	7	40	0	47	53
8. *Met. at side of display: not visible*, rest as for 1.	6	39	6	22	0	67	33
6. *Axis of display vertical*, rest as for 1.	10	10	13	44	0	67	33
4. *Neon lights*, 1 dist. apart, rest as for 1.	11	45	0	55	0	100	0
5(A). *Interrupted rhythm* (1·5 *sec.*), 2 dists. apart, cams for met.	10	15	10	30	0	55	45
5(B). *Interrupted rhythm* (3 *sec.*), rest as for 5A.	Same as for 5A	5	35	25	0	65	35
10. *Interrupted rhythm* (6 *sec.*), rest as for 5A.	5	0	0	10	0	10	90

V

DISCUSSION

Two main points emerge from our second group of experiments when these are considered together with the results previously reported by us. The first, and subsidiary, point is the very wide range of variation in the total number of movement perceptions of any kind observed under different experimental conditions. The

second point is that with which we have been primarily concerned, namely the factors affecting the path which the apparent movement is taking.

1. *Facilitation and inhibition of apparent movement.*

In the main experimental conditions, where three inter-light distances were used, the range of variation of total movement, expressed as a percentage of the total number of possible responses, is from 93 in Condition 1, to 47 in Condition 7. In other conditions, where only one, or two, inter-light distances were used, the range is from 100 in Condition 4 to 10 in Condition 10. The results for Condition 1 seem to indicate a definite facilitation in the perception of movement of any kind, as, even in comparison with Condition 2, where the metronome was inaudible to the subjects but where otherwise the experimental situation was identical, the proportion of total movement is significantly higher ($\chi^2 = 5{\cdot}45$, $P = < 0{\cdot}02$).

This seems to confirm Werner and Zietz's (op. cit.) finding of a facilitation of movement perception under certain conditions of auditory accompaniment, while our two additional findings with Condition 7 (metronome visible and audible, at side of display) and Condition 8 (metronome audible but not visible, at side of display) indicate that an accentuation of the background stimuli, visually or auditorily, or both in conjunction, can result in a relative inhibition of movement perception. The common factor for these two conditions would seem to be the displacement of the metronome to the side of the display screen, thus perhaps dissociating the auditory rhythm from the visual stimuli. Particularly in Condition 7, the actual showing of the metronome in action to the side might reasonably be considered to have "broken-up" the strong movement pattern of the original Condition 1.

Explanation of these effects is clearly closely bound up with the factors producing the pendulum phenomenon itself, particularly the factor of past experience, and we shall leave further discussion of this to the following section.

2. *The arcuate path of apparent movement.*

The second main point apparent from the summary of results in Table X is the difference in the quality of the perceived movement within experimental conditions which have resulted in an almost identical proportion of total movement responses. Thus, the percentage of total movement in Conditions 2, 3, 6, 7 and 8 varies only from 67 to 77, but the percentage of vertical curve responses varies from 3 in Condition 3 to 50 in Condition 7. The range of vertical curve for all the main conditions is from 0 in Condition 10 to 77 in Condition 1.

We shall now consider in detail three possible factors which, either separately or in combination, may account for these differences in the path of movement.

(1) *Past experience.*

Evidence supporting the effects of experience on the perception of apparent movement may be, for our purposes, classified into two kinds:

(*a*) Expectancy, and familiarity with the content of the display—for example, Piéron's (1934) experiments, using realistic objects familiarly expected to move, and Wertheimer's (op. cit.) study of experimentally-induced set.

(*b*) Cortical elimination experiments, such as those of Smith (1948) and others, where the images of the two stimulus objects are supposed to be separately projected so that such movement as is seen is the product of part-movements, retinally produced, and a central interpretation given to them.

In our first experiment, it seemed possible to us that past experience might be acting to modify the path of apparent movement owing to certain conditions of the display. The display used in Condition 1 had a number of elements somewhat similar to the stimulus-pattern of a mechanical pendulum system, particularly a clock pendulum. A bright disc appeared to move from side to side, in a regular, rhythmical fashion, accompanied by the audible ticking of a metronome. It seemed possible therefore that memory traces derived from everyday experience of clock pendula were being activated by these conditions, and were in some way serving as a background to the observer's immediate perceptual responses. Such a background may be adduced either as the result of a deliberate "effort after meaning," or as the result of a much less explicit matching process. Whichever were the case, the effect might be to modify the movement perception to conform to that which a real pendulum would be expected to take.

If this "experiential" hypothesis is correct, then it should follow that the greater the resemblance of the display conditions to the stimulus pattern of a clock pendulum in oscillation, the greater will be the likelihood of perceived movement taking a pendular form. Reference to Table X shows how far this is borne out by the experimental data. In Condition 1, the resemblance of the stimulus conditions to that of an actual clock-pendulum is certainly greatest, with the regular rhythm, audible metronome etc. It is in this condition that the incidence of "vertical curve" (i.e. pendular) movement is greatest. All other experimental conditions involve a diminution in the degree of resemblance to an actual clock pendulum. For instance, in Condition 2 there is no audible ticking, in Conditions 3 and 9 the lights are square, in Conditions 5 and 10 the rhythm is irregular, and in Condition 6 the axis of display is vertical. Under all these other conditions, a reduction in pendular movement perception takes place in comparison with Condition 1.

Regarded in this way, the data may be said to provide some evidence in favour of the experiential model hypothesis. Further confirmation comes from subjects' reports, which frequently imply the operation of a background model, and from the fact that, when subjects observe two different experimental conditions consecutively, they tend to persist in reporting their initial type of movement perception under the second condition, whereas new subjects would report a different type of movement.

The "model" hypothesis fails, however, to explain the following facts:

(i) Condition 9 (square lights at 45°) produces far more pendular movement than Condition 3 (square lights parallel to the main axis of display), although it could hardly be said that one condition approximated more closely to the real model than did the other.

(ii) The occurrence of horizontally-curved movement.

(iii) The inhibitory effect of neon lamps (Condition 4).

(2) *Physiological nystagmus.*

Two types of experimental evidence have been brought forward in previous work to disprove the hypothesis that following eye-movements play an important part in the perception of apparent movement. These are:

(*a*) Cine-photography gives no evidence of eye-movements consistently following the stimuli.

(*b*) Apparent movement can be perceived in two opposite directions at once.

Objections on these grounds are justly applied in respect of coarse following and fixating movements of the eye, but it seems possible that the much finer and more

rapid trembling movements classified as physiological nystagmus may at least contribute to the directing of the path of apparent movement in our own experiments.

The movements comprising "physiological nystagmus" are described as being continuous, oscillatory and involuntary (Duke-Elder, 1932), and they are said to occur with such rapidity and fineness that they may be undetectable by ordinary cine-camera techniques. It would therefore be possible for several nystagmic movements to occur in each direction in the time-interval between the two stimuli of an apparent movement display such as we have used.

The relevance of this hypothesis as a part-explanation of arcuate movement may be understood from the fact that Duke-Elder (op. cit.) and De Jong (1946) distinguish two kinds of nystagmic movements. These are:

(i) Jerky nystagmus, in which a slow movement in one direction is followed by a rapid return to the original position;

(ii) Pendular nystagmus, in which the movements from side to side are smooth, undulatory, and of equal speed in either direction.

According to Duke-Elder, a type of pendular nystagmus is found in children reared in dark surroundings, and can normally be seen in the cat when illumination is dim.

It seems possible, therefore, to the present writers that two features of our experimental situation, namely the regularity of the rhythm with which the stimuli appeared in most conditions, and the low level of general illumination, may have been conducive to the occurrence of a pendular nystagmus, and may have contributed to the perception of pendular apparent movement. Such a hypothesis is supported by the fact that, where the rhythm was irregular (Conditions 5 and 10), pendular movement was relatively rare. It fails, however, to explain why a regular rhythm should not have induced much arcuate movement when the stimulus lights were square.

(3) *Intervening adaptation.*

We have seen that neither of the two factors so far discussed have been adequate to account for all the facts in our two series of experiments. It is proposed, therefore, to review a third explanation derived from the so-called "statistical" theory of visual acuity as propounded by Marshall and Talbot (1942) and applied to the explanation of figural after-effects by Osgood and Heyer (1952). Such a theory seems in many ways preferable to the Gestalt conceptions of "physiological short-circuits," leakage of excitation across the cortex, and the operation of extra-neural field-forces.

Marshall and Talbot drew attention to certain physiological mechanisms whereby a pattern of visual stimuli can become both spatially extended and temporally prolonged on its route from the retina to the visual cortex. These mechanisms include diffraction in the pupil, physiological nystagmus, reverberatory activity, multiplication of unit pathways, and reciprocal overlap. The authors conclude that the cortical representation of a contour is probably a normal distribution of excitation symmetrical about its axis. If this is correct, it would follow that the cortical representation of the stimulus shapes used in our experiments would consist of a region of relatively intense and homogeneous excitation surrounded by a region in which the intensity of excitation falls off towards the periphery. Thus the two stimuli will be represented by two foci of activity in the cortex which have no clear-cut boundaries. Under certain conditions of stimulation, considerable summation might be expected to take place in the region between the two foci, and this could furnish the excitatory basis for the perceived path of apparent movement.

Such an account would lead one to expect the apparent movement to take place always along the most direct path between the two stimuli, but it seems possible to extend the implications of the statistical theory to our own pendulum phenomenon by bringing in the factor of satiation or adaptation. Sanders (1921) is reported by Bender and Teuber (1946) to have shown that prolonged stimulation of a retinal area inhibits stroboscopic motion across it. Detherage and Bitterman (1952) made a similar observation, and also noted that, when the two stimuli of an apparent movement display are placed slightly to one side of an area occupied immediately beforehand by a circular inspection-figure, the path of apparent motion is reported as being curved concentrically with the previously displayed circle.

These observations no doubt arise from a set of underlying cortical processes similar to those producing figural after-effects in general. They are important to our experiments because they suggest one possible condition within the visual projection system—i.e. a region of localized adaptation between the two postulated foci of excitation—which might account for the deviations in the path of apparent movement reported in some of our experimental situations.

A possible explanation in terms of such an "intervening adaptation" is outlined below, and is schematically represented in Figure 1, where the three different light-shapes used in our experiments are shown.

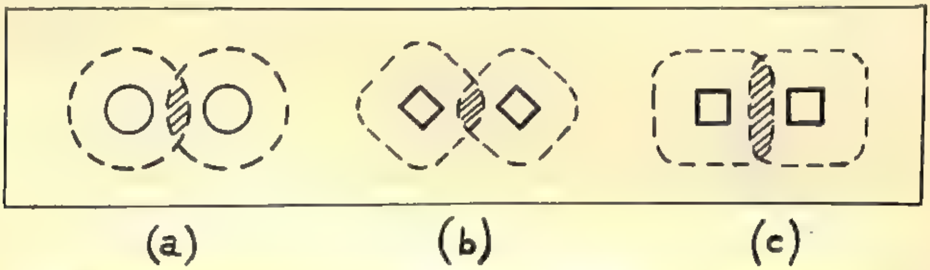


FIG. 1. Diagram of hypothetical regions of adaptation for three different light-shapes.

In Figure 1, the inner figures (solid lines) represent the supposed spatial arrangement in the visual cortex of the two foci of excitation. The area between these foci and the outer broken lines represents a region in which some peripheral excitation might be expected to occur. Where the two peripheral regions overlap (shaded area), some summation of activity from either focus should take place. Considering a single pair of stimuli, or recurrent pairs separated by an appreciable interval of time, the path of apparent movement ought to be straight, since the region of greatest summation would be that directly intermediate between the two foci.

Where the rhythm of presentation is regular, however, the situation may be altered, since repeated excitation of the region of overlap from foci on either side of it will mean that cells in this region are firing in bursts which occur twice as frequently as those in either focus considered by itself. Given a critical relationship between stimulus-frequency and recovery-cycle, this may cause the region of overlap to become progressively adapted, so that after a number of exposures the cells in this region will cease to respond to stimulation. A gradient of relatively more intense excitation would then be created round the periphery of the region of overlap, and this gradient might then form an excitatory basis for the path of apparent movement.

Such an account would seem to fit the three situations represented in Figure 1. In (*a*) the region of peripheral overlap is fairly compact and circumscribed. This is even more the case in (*b*) with the diamond-shaped lights, but in (*c*) the region of

overlap is narrow and elongated. As the region of adaptation corresponds to the region of overlap, this seems to offer a reasonable explanation for the fact that the highest proportions of arcuate movement are reported in Conditions 1 and 9 (77 and 61 per cent. respectively), as against 3 per cent. in Condition 3 where the stimuli were square; square stimuli, tending to produce a long and narrow region of adaptation, should lead to the perception of straight movement with the middle part of the path being seen as indistinct.

It can be seen, therefore, that this intervening adaptation hypothesis may adequately account for the following facts:

(1) The greater incidence of curved movement when the interval between pairs of stimuli is reduced (less recovery from adaptation being possible).
(2) The greater incidence of curved movement when the distance between stimuli is reduced (the region of peripheral overlap being greater).
(3) The greater incidence of curved movement when the lights are circular or diamond-shaped, as compared with that observed with square lights.

It does not, however, explain why altering the axis of display from horizontal to vertical should so reduce pendular movement, and it fails to account for certain minor variations, such as the effects of making the metronome inaudible, using neon lights, transfer situations, and so on. On the basis of such a theory also, it might have been predicted that, when the original display situation with circular lights, etc., was presented continuously for about 15 minutes (Condition 11), perception of movement would begin as straight, then change to pendular, and remain fairly consistently pendular thereafter. Our evidence under this Condition is not at all conclusive, and this situation might well be repeated. Finally, it is difficult to understand, on the adaptation hypothesis alone, why the curve should have generally been reported as dipping rather than arching.

Conclusions

We emphasized initially that no single explanation, but a combination of explanations, might be required to account for the various apparent movement phenomena observed in our experiments. This was to be expected from reviews of visual apparent movement experiments such as that of Neff (1936), whose conclusion was: "Of the twenty or more significant variables which have now received treatment, we can simply make the general statement that any or all of them, in a large number of possible forms of combination, may suffice to bring about a perception of movement" (p. 37).

Our conclusion must be that each of the three hypotheses we have discussed is necessary to explain one or more of the sets of observations in our experiments. It seems reasonable to suppose that the hypothesis of intervening adaptation can account primarily for the effects of light-shape variation under conditions of rhythmical presentation. By itself, however, it does not account for the difference brought about by presenting the circular lights on a vertical axis. Here, Osgood and Heyer's (op. cit.) account of the contributory effect of physiological nystagmus may be directly relevant, for they point out that: "Owing to the manner in which the ocular muscles are attached and balanced (and probably also to some extent owing to experience), these fine eye movements are more extensive in the horizontal than in the vertical plane" (p. 110). This provides a further link between physiological and experiential explanations, and may also account for the fact that, usually, pendular movement is perceived as dipping down and not arching over the horizontal axis of display.

Finally, there is little doubt that experience of certain familiar models, such as clocks, pendula, etc., may be a contributory factor, where the conditions of stimulations are, as it were, finely balanced between the possible perception of one form of movement and another. Further, as Osgood and Heyer point out, the way in which physiological nystagmus takes place may be partly a product of experience. Our results, however, could certainly not support an uncompromising experiential theory such as Piéron's (1934, 1952), where he states that, if physiological systems intervene in movement perception, they do so only in so far as they have been ". . . modelled by the repetition of perceptual experiences along the lines of associative mechanisms revealed by the study of conditioned reflexes" (1934, p. 226).

References

De Jong, R. N. (1946). Nystagmus: an appraisal and a classification. *Arch. Neurol. Psychiat.*, **55**, 43–56.

Detherage, B. H., and Bitterman, M. E. (1952). The effect of satiation on stroboscopic movement. *Amer. J. Psychol.*, **65**, 109–18.

Duke-Elder, W. S. (1932). *Text Book of Ophthalmology*, Vol. 1, London.

Hall, K. R. L., Earle, A. E., and Crookes, T. G. (1952). A pendulum phenomenon in the visual perception of apparent movement. *Quart. J. exp. Psychol.*, **4**, 109–20.

Marshall, W. H., and Talbot, S. A. (1942). Recent evidence for neural mechanisms in vision leading to a general theory of sensory acuity. In: H. Klüver, ed. Visual mechanisms. *Biol. Sympos.*, **7**, 117–64.

Neff, W. S. (1936). A critical investigation of the visual apprehension of movement. *Amer. J. Psychol.*, **48**, 1–42.

Osgood, C. E., and Heyer, A. W. (1952). A new interpretation of figural after-effects. *Psychol. Rev.*, **59**, 98–118.

Piéron, H. (1934). Remarques sur la perception du mouvement apparent. *Année Psychol.*, **34**, 245–48.

Piéron, H. (1952). *The Sensations.* London.

Sanders, E H (1921) Nederl. tijschr. v. geneesk., 65, 1820–1836: quoted by Bender, M. B., and Teuber, H. L. (1946). *Arch. Neurol. Psychiat.*, **55**, 627–58.

Smith, K. R. (1948). Visual apparent movement in the absence of neural interaction. *Amer. J. Psychol.*, **61**, 73–8.

Vernon, M. D. (1952). *A Further Study of Visual Perception.* Cambridge.

Werner, H., and Zietz, K. (1928). Über die dynamische struktur der Bewegung. *Zsch. f. Psychol.*, **105**, 226–49.

Wertheimer, M. (1912). Experimentelle Studien über das Sehen von Bewegung. *Zsch. f. Psychol.*, **61**, 161–265: quoted by Vernon, M. D. (op cit.).

APPARENT TRANSPARENCY AND THE TUNNEL EFFECT

BY

ALAN J. GLYNN, O.F.M.

From the Laboratory of Experimental Psychology, Louvain.

SUMMARY

Experimental evidence can be adduced to show that when an object passes behind an opaque screen, its form may appear unchanged even though certain of its parts are invisible ("screen effect," "tunnel effect"). It has also been reported that, under certain conditions, the form continuity of the object can be directly seen because the opaque screen appears to be transparent (Rosenbach phenomenon). The aim of the present research was to clarify the relationship between these phenomena.

The experimental method consisted in presenting to the observers a grey rectangular figure (12 cm. × 1·2 cm.), moving across the ground glass of a camera, to which an opaque screen (15 cm. × 1 cm.) of lower brightness was attached.

It was found that the apparent transparency of the screen depended, in part, on the subject's preparation for the task. In addition to the factor of "set" practice was also important.

The analysis of the responses of the observers showed that the form continuity of the figure appeared under three distinct, but related, forms of structural organization: an unshaped darkening on the screen, a shaped darkening on the screen, and transparency. Each type of response could be favoured by an appropriate instruction. Furthermore, any change in the stimulus situation affected them all in the same way.

The other factors conditioning the form continuity of the figure pertained to the stimulus situation. These were the speed of the translation motion of the figure, the dimensions of the figure and the screen, and the brightness relation between them.

The study of the Rosenbach phenomenon showed that it was not related to the "screen effect" and the "tunnel effect."

Il a été démontré expérimentalement que la forme et la grandeur apparentes d'un objet peuvent demeurer inaltérées bien que certaines de ses parties soient invisibles ("effet écran," "effet tunnel"). On a signalé d'autre part que, dans certaines conditions, un écran opaque qui recouvre partiellement un objet peut sembler transparent et qu'on a alors l'impression de voir l'objet dans sa totalité (phénomène de Rosenbach). Le but de la présente étude était de préciser les relations existant entre ces phénomènes.

La méthode expérimentale consistait à montrer aux observateurs l'image d'une figure rectangulaire de teinte grise (12 cm. × 1·2 cm.) projetée sur le verre dépoli d'une chambre photographique et parcourant celui-ci de part en part, tout en croisant un écran opaque plus sombre (15 cm. × 1 cm.) fixé au milieu de la face antérieure du dit verre dépoli.

On a trouvé que la transparence apparente de l'écran dépendait en partie de la préparation du sujet, de son attitude mentale et de l'exercice acquis.

L'analyse des réponses des observateurs a montré que la continuité de la forme de la figure pouvait se présenter de différentes manières: comme obscurcissement local informe de l'écran au moment du passage de la figure; comme obscurcissement de forme définie continuant le contour de la figure; et enfin sous l'aspect de la transparence. Chaque type de réponse peut être favorisé par une instruction appropriée, et, de plus, des changements apportés au système de stimulations les affecte tous dans le même sens.

Les autres facteurs déterminant la continuité apparente de la forme ressortissent au système de stimulations: la vitesse de translation de la figure, ses dimensions et celles de l'écran, et surtout les rapports de brillance des différentes parties du champ.

L'étude du phénomène de Rosenbach a permis d'établir qu'il était d'un autre ordre que "l'effet écran" et que "l'effet tunnel".

I

INTRODUCTION

A series of experiments by Michotte and his collaborators (Sampaio, 1943; Knops, 1947; Michotte, 1950) has shown that if the length of a moving object (e.g. a rectangle) is gradually decreased after it has attained the border of a stationary

object (e.g. another rectangle), the former does not appear to the observer as diminishing progressively in size, but as retaining its initial dimensions and "sliding behind" the latter ("screen effect"). Similarly, under appropriate conditions, the movement of an object passing behind a screen may appear absolutely continuous, even though one of its phases is hidden by the screen ("tunnel effect"); *vide* Burke, 1952. Michotte has coined the expression "amodal aspect of perceptual experience," to designate the phenomenal presence of the hidden part of an object or process when there are no sensory modalities corresponding to that part.

Many years ago, Rosenbach (1902) reported that, under experimental conditions more or less similar to those of the "screen effect" and the "tunnel effect," the observers actually saw the form and colour of the parts of an object, through the opaque screen which covered them. That is, in this case, there was a "modal presence" of these parts as the screen seemed to be transparent. The "modal presence" of the covered parts in Rosenbach's experiment and their "amodal presence" in the "screen effect" and "tunnel effect" suggested that the two might differ only in degree and not in kind. If such a continuity could be established, the "amodal presence" of the covered parts in the latter cases could eventually be considered as the lowest limit of the apparent transparency, corresponding to a system of excitation inferior to its threshold. The aim of the present research was to study the conditions of the Rosenbach phenomenon systematically in order to ascertain if such a hypothesis is tenable.

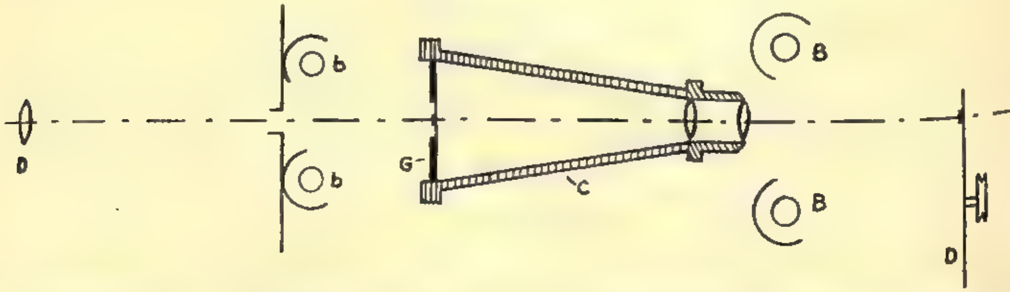
Rosenbach's experimental material consisted of coloured figures of various shapes (triangles, ovals, rectangles), placed behind a narrow band (1 cm.) of black opaque paper in such a way that parts of the figure protruded on either side of it. He instructed his observers to move the figure back and forth beneath the stripe and then asked them if it was transparent. Practically all the observers affirmed that it had to be, since they saw the figure through it. They saw a mixture of the colour of the figure and of the stripe on the crossing-point.

II

Method

The procedure eventually adopted for use in the research, evolved out of a series of preliminary trials in which other conditions were used and then rejected. Our aim was to reproduce as striking an impression of transparency as possible within the framework of Rosenbach's conditions, but in such a way that each factor of the objective situation could be controlled for systematic study. As a result of these trials, two modifications were introduced into Rosenbach's original procedure. The first was the use of a projection device (a photographic camera). This enabled us to eliminate disturbing shadows

FIGURE 1



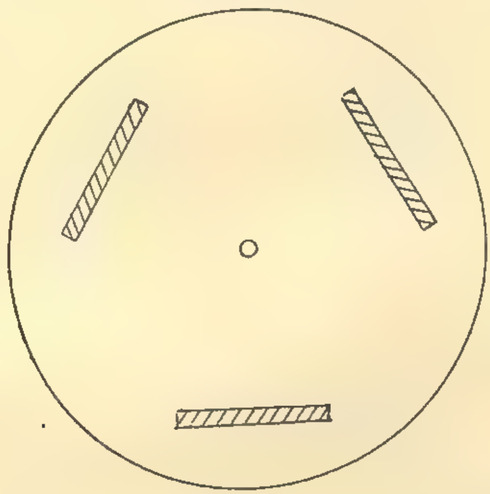
O = Observer. b = Small bulbs. G = Ground glass with shades and, in the middle, the opaque screen. C = Photographic camera. B = Large bulbs. D = Rotating disc.

cast by the screen on the background when the situation was viewed directly. In addition, it allowed systematic and controlled variations of all sorts. The second modification was the limitation of the study to achromatic material.

A schematic drawing of the experimental arrangement as it was set up in the dark room of the laboratory appears in Figure 1. The subject viewed the ground glass of the camera through an opening (5·7 cm. × 6·5 cm.) in the black screen, with his head firmly fixed in the headrest. A vertical stripe (15 cm. × 1 cm.) of opaque paper, placed at the centre of the ground glass, served as the screen. Two adjustable flaps limited the visible area of the ground glass on either side of it to 1·5 cm. This distance between the border of the screen and the border of the flaps was kept constant even when wider screens were placed on the ground glass. The two small bulbs, attached to the reverse side of the observer's screen, provided an illumination on the screen and the ground glass.

A disc of white cardboard, bearing three grey rectangular figures set tangentially to its centre, rotated at a constant speed in front of the camera. (Fig. 2.)

FIGURE 2



The image which the disc projected on the ground glass, was that of a rectangle (11 cm. × 1 cm.) on a white background, moving uniformly (linear speed: 18 cm. per sec.) in an arc of a circle across the field. At the point of its trajectory where the figure lay across the middle of the screen, it appeared to the observer to pivot suddenly about an apparent horizontal axis, that is, about an axis perpendicular to the surface of the screen. (Fig. 3.)

The standard conditions of brightness of the opaque screen, figure, and background were as follows:

The reflection factor of the "normal" opaque screen amounted to 0·34, and its illumination provided by the small bulbs was ± 35 Lux.

Taking this brightness as unity, the relative brightness of the moving rectangle (figure) was 1·6 and that of the white background 2·2, the transmitted light through the ground glass and the reflected light on its surface being both taken into account.

This method was used throughout the experiments except where a change in the experimental procedure is explicitly indicated.

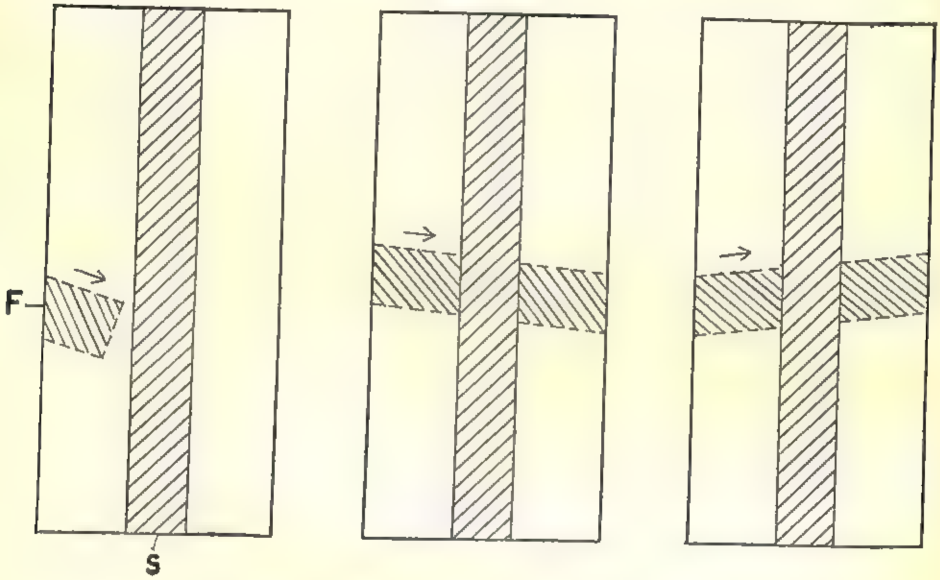
III

The Preparation of the Subject

In his study of the "tunnel effect," Burke, 1952, found that it was sufficient to ask untrained observers what they saw in order to obtain an accurate description of the movement of the object, its continuity, discontinuity, etc. It soon became clear that, even under our optimum physical conditions, a neutral instruction was insufficient to obtain a report on the phenomenon which we intended to study. In an experiment with such an instruction under standard conditions, only 2 subjects in a group of 25 untrained

subjects reported that, on one of the three presentations of the figure passing the screen, they thought they saw a shadow on it at the crossing-point. The other 23 subjects limited themselves to a detailed account of the appearance of the figure, the course of its lateral movement across the ground glass, and the character of its swinging movement as it passed the screen. The subjects were then given three more presentations and asked two further questions: whether a part of the figure was hidden by the screen at the crossing-point and how they would describe the screen in relation to their response. Twelve subjects (48 per cent.) reported seeing the figure as continuous in form, 9 subjects (36 per cent.) saw it as hidden by the screen (i.e. as a simple "tunnel effect"), and 4 subjects (16 per cent.) spoke of a "shadow" at the point of crossing. Of the 12 subjects who reported the figure as continuous, only 1 subject referred to the screen as "vaguely transparent." With repeated trials, all save 2 of the 25 subjects reported that the rectangle was uninterrupted in form as it passed the screen. Only 15 subjects, however,

FIGURE 3



S = Opaque screen placed on the ground glass.
F = Figure (moving rectangle)

(both screen and figure were under standard conditions plain greys of different brightness—see text)

described the screen as transparent. The remaining 8 subjects reported that the screen was not transparent, but appeared to be darkened on the area where the rectangle passed behind it.

Two other systematic experiments were undertaken to clarify certain results of this initial experiment. The first of these was devised to determine the role of practice on the report of figural continuity under an instruction to note whether the contours of the figure were continuous or interrupted by the screen. Sixteen untrained subjects took part in this experiment. The second experiment compared the responses of thirty untrained subjects under two different instructions. Approximately half of the group were asked first about a brightness change on the screen and then about its transparency. The rest of the group reported first on the transparency of the screen and then on the brightness change. Both experiments were carried out under standard conditions except for the screen which was only 30 mm. high. The responses of the subjects were recorded after each of three presentations in the first experiment. In the second experiment, the subjects gave their reports after three presentations.

The observers' reports on these three experiments led to certain significant findings on the influence of "set" and learning on the structural organization of the stimulus field. It is appropriate to state them as general conclusions before the analysis of the objective

conditions of the Rosenbach phenomenon, since the latter are only effective if these factors are present.

1. The response of figural continuity was, at least in part, the outcome of a certain "set." Unless the observers were induced, through the instructions, to concentrate their attention on the form completion of the figure, they did not mention this aspect of the field, but reported only on the movement of the figure, its direction, speed and uniformity. The situation thus appeared to them initially as a pure "tunnel effect." Our results indicate that this detail was omitted not because the observers considered it unimportant, but because they did not notice it at all. Even with repeated presentations of the figure passing the screen, there was no spontaneous mention of the contours of the figure as continuous under a neutral instruction. This finding fully substantiated Rosenbach's (1902) claim that the form continuity of the figure was, in part, a function of the instruction which induced the observer to adopt a certain attitude *vis-à-vis* the stimulus field.

2. A second factor helping to elicit the response that the figure was continuous, was that of learning. When the observers were instructed to report whether the contours of the figure were continuous or hidden by the screen, most of them reported only a darkening of ill-defined shape and indefinite localization (a reversal of simultaneous brightness contrast) on a single presentation of the experiment. With additional practice, (two presentations were sufficient for the majority of the subjects in the second experiment), the brightness change assumed both of these characteristics so that the figure was described as continuous in form. The manner in which the darkening was localized divided these responses into types. The first type of response was of a precise darkening *on* the screen. This corresponded exactly to the form of the figure so that its contours were clearly delineated on the screen. The subjects found it extremely difficult to define the material character of the screen in relation to this darkening effect. Either they said that it was not transparent or they replied that they did not know whether it was transparent or not. In the second type of response, the subjects reported that they saw the figure as "dark" *through* the screen. They described the latter as transparent either directly or by an equivalent expression (eventually translucent). The two characteristics which distinguish this type of response from the preceding were: (1) the brightness change as assigned to the figure and located in its place; and (2) there was a phenomenal separation of the retinal stimulation corresponding to the crossing-point so that both figure and screen were represented as having a distinct brightness of their own. This phenomenal dissociation of the retinal stimulation into two components, one of which is appropriated to one object and the other to the second seen through it, is the salient characteristic of the impresssion of transparency.

It is to be noted that, although the observers defined the screen as transparent, they did not report that the covered part of the figure appeared in its true colour. Highly trained observers occasionally gave such reports.

3. Finally, these experiments showed that the three responses, unshaped darkening, shaped darkening, and transparency were related. The latter responses arose as further specifications of the first. A comparison of the frequency of the responses of a shaped darkening and of transparency revealed that, when the experimental situation was presented to the same observers under an instruction equally favourable to each type of response, the first type of figural continuity was more easily attainable than the second.

IV

The Stimulus Pattern

The influence of "set" and learning on the form completion of the figure obviously depended on determined stimulus conditions. This led to a second series of experiments in which the various factors of the visual field were systematically modified in order to determine the optimum physical conditions for the Rosenbach phenomenon. The subjects who took part in these new experiments were 5 members of the laboratory staff (Mi, Mo, Kn, Ge, Gl) and 19 of the 25 subjects who had participated in Experiment 1. Of the laboratory personnel, two (Mo, Ge) consistently saw the continuity of the figure as a shaped darkening on the screen. The three others (Mi, Kn, Gl) generally gave a transparency response, but, under less favourable conditions, located

the brightness change on the screen. The 19 subjects were similarly divided on the type of response which they gave: 14 subjects (74 per cent.) described the screen as transparent; the remaining 5 subjects (26 per cent.) located the dark area on the screen. The predominance of transparency responses among the 19 subjects was due undoubtedly to the suggestion given them by the third instruction of the first experiment. Regardless of the particular form of response, the stimulus factors which were conducive or detrimental to seeing the figural completion were the same in both cases.

The form completion as the result of the extension of the contrast relation between figure and background to the screen.

The two types of response which implied the form completion of the figure possessed a common element, a darkening, at the crossing-point of the screen. What was singular about this brightness change was that it was exactly the reverse of the contrast effect which could be expected on the area of the screen lying between the projecting parts of the figure. Instead of making this area of the screen brighter than its other parts, the projecting parts of the figure acted in the opposite sense. Two experiments were performed to determine what this inverse induction represented.

In the first of these experiments (Experiment 4) a black background (relative brightness 0·6—see page 127) was substituted for the usual white background of the standard figure. The four members of the staff (Mi, Kn, Ge, Gl), who served as subjects, were given several presentations of the experimental situation. All the subjects reported that the completion now appeared as a well-defined bright area and not as a dark area. A similar brightness reversal took place on the parts of the screen adjacent to the crossing point: they appeared darker than the critical area while, under standard conditions, they were always reported as brighter than it.

The reports of this experiment led to the conclusion that the primitive fact in the form completion of the figure was the maintenance of the contrast relation between the figure and the background across the screen.

The second experiment (Experiment 5) was designed to test this conclusion by the simple expedient of suppressing the surface contrast relation and noting its effect on the form continuity of the figure. Mere outlines of the rectangular figures, drawn with black ink on white backgrounds, were presented to four members of the staff. Only 1 subject (Ge) reported the contours of the rectangle as continuous on the first trial after several presentations. Another subject (Mi), who had tried many times previously without success, saw the designed figure as continuous after the first subject (Ge) reported, in his presence, the contours as faintly continuous. The third subject (Kn) saw no continuity on his first trial after more than twenty presentations. On the following day after an equivalent number of presentations, he finally had some impression of continuity. The final subject (Mo) never reported the contours of the figure as continuous on three trials made on three successive days.

The data derived from these two experiments completed one another: the first links the change on the screen to the contrast between the brightness of the surface of the figure and that of the background; the second shows that once this factor is suppressed, the figural continuity practically disappears. The apparent change on the screen may thus be described as an extension of the contrast situation existing between the figure and the background to the screen, or as a mixture of the screen's brightness both with the brightness of the figure and that of the background. In his account of the phenomenon, Rosenbach only considered the mixture between the figure and the screen. These experiments show that a brightness change takes place not only at the crossing-point, but also on the parts of the screen immediately

adjacent to it. This area, which extends a few mm. above and below the crossing-point, appears brighter than it with figures set on white backgrounds and darker than it with figures set on black backgrounds.

The influence of the brightness relation between the figure and the screen.

Under optimum stimulus conditions, there was only a slight brightness difference (low contrast) between the figures and the screen. A systematic experiment was designed to study the degree of contrast difference consistent with the form continuity of the figure.

Experiment 6.

Experimental conditions.—The figure was the standard gray. Four screens of different brightness, with a reflection factor of 0·75, 0·34 0·19 and 0·04 respectively, were presented on the ground glass in pairs. The visible area of the ground glass was increased to 10 cm. in width and the screens placed 2 cm. from its centre on either side.

Subjects and procedure.—The subjects were 19 of those who had previously served in Experiment 1. Two screens of different brightnesses were shown simultaneously on the ground glass. The screens were then removed and replaced by two others until each screen had been compared with all the others. Each subject made two such comparisons for each combination of screens. This order as well as that of the position of the brighter screen on the ground glass, was reversed in the second series. The experimenter instructed the subjects to look first at the screen on the right side of the ground glass and then at the one on the left. Two presentations were given for each screen. The subject was asked to report whether the figure was continuous in both cases and which screen gave the better impression of continuity.

Results.

Table I summarizes the results of this experiment. A chi-square analysis (expected frequencies for all brightnesses: (321) (114/456) shows that the frequencies are significantly different ($\chi^2 = 53{\cdot}5$) at the 0·001 per cent. level for 3 d.f.

TABLE I

FREQUENCIES OF FORM COMPLETION RESPONSES WITH INCREASING BRIGHTNESS DIFFERENCES BETWEEN FIGURE AND SCREEN

*Screen brightness**	*N*	*Absolute frequency*	*Per cent.*
0·75	114	24	21
0·34	114	106	93
0·19	114	96	84
0·04	114	95	83
	456	321	

* Factors of reflection.

If the brightest screen is excluded from the Table, the χ^2 is equal to 0·6 which is not significant for 2 d.f. The reason for the difference in the two cases is evidently connected with the brightness relation between the figure and the screen. The white screen appeared much brighter than the figure. With all the other screens, the brightness gradient went in the opposite direction: the screen always appeared darker than the figure. Where this relation was observed, there were no significant differences in the frequencies of the completion responses. Thus screens of 0·19 and 0·04 brightness

were also consistent with figural continuity. The completion, however, was not of equal impressiveness. The standard screen was the most favourable for that response. The white screen, on the other hand, was most unfavourable. The loss of figural

TABLE II

RELATIVE FREQUENCIES WITH WHICH THE SCREEN GIVEN AT THE TOP WAS PREFERRED TO THOSE AT THE LEFT

*Screen**	0·75	0·34	0·19	0·04
0·75	—	0·87	0·74	0·71
0·34	0·05	—	0·26	0·21
0·19	0·05	0·63	—	0·87
0·04	0·13	0·68	0·53	—

* Factors of reflection.

continuity with this screen may have been due to a strong contrast effect between the figure and the screen counteracting the opposite induction (darkening).

A second method for studying the need of a determined direction of the brightness gradient between figure and screen was to leave the screen constant and vary the brightness of the figure.

Experiment 7.

Experimental conditions.—Figures of four different brightnesses (see Table III) were shown on white backgrounds. As they appeared on the ground glass, the first two were brighter than the stand ard screen, the last two darker.

Subjects and procedure Five members of the staff (Mi, Mo, Kn, Ge, Gl) served as subjects. Two presentations of the different figures were given. Four observations were made by each subject on different days beginning alternately with the two extremes of brightness.

Results.

The results of Experiment 7 are given in Table III. No statistical analysis is needed to reveal the detrimental effect of changing the direction of the brightness gradient in the case of figures of different brightnesses seen on white backgrounds.

TABLE III

FREQUENCIES OF FORM COMPLETION RESPONSES WITH INCREASING BRIGHTNESS DIFFERENCES BETWEEN THE FIGURE AND THE STANDARD SCREEN

*Figural brightness**	*N*	*Absolute frequency*	*Per cent.*
1·6	20	20	100
1·25	20	20	100
0·8	20	2	10
0·6	20	0	0

* Relative brightness values. See p. 127.

As a further check on the influence of the direction of the brightness gradient between figure and screen, the 0·6 figure on a white background was tested with a black screen (Experiment 8). This restored the proper brightness relation. The three

subjects (Mi, Ge, Gl) unhesitatingly reported the figure as completed on the first trial. This excluded the possibility of ascribing the negative results with this figure in Experiment 7 to the brightness difference between the figure and the background.

The same subjects served in an additional experiment with the figures of different brightnesses on *black* backgrounds. Its purpose was to determine whether the figural completion demanded a brightness gradient which was the reverse of that required for the same figures seen on white backgrounds (Experiment 9). That is, did the screen have to be brighter than the figure when the latter was presented on a black background? In carrying out this experiment we met with technical difficulties due to the light reflected on the ground glass. For this reason we will limit ourselves to saying that we found some indications that this was the case.

The influence of the width of the screen.

The importance of the brightness value of the screen for the continuity of the figure suggested that its width might also exercise some control on this response. This was investigated in the following experiment.

Experiment 10.

Experimental conditions. Screens of standard brightness with the following widths: 5 mm., 10 mm., 20 mm., 30 mm., 40 mm., 50 mm., and 60 mm. were attached successively to the ground glass.

Subjects and procedure. Five members of the staff (Mi, Mo, Kn, Ge, Gl) served as subjects. Four observations were made by each subject on different days beginning alternately with the screens of smallest and largest width.

Results.

The results are presented in Table IV. No statistical analysis is required to deduce the conclusion that the figural completion can only extend a certain distance.

TABLE IV

FREQUENCIES OF FORM COMPLETION RESPONSES WITH SCREENS OF INCREASING WIDTH

Width of the Screen (mm.)	*N*	*Absolute frequency*	*Per cent.*
5	20	20	100
10	20	20	100
20	20	9	45
30	20	3	15
40	20	0	0
50	20	0	0
60	20	0	0

The influence of the phenomenal appearance of the screen.

Additional properties of the screen were investigated in connection with the fact, discovered early in the research, that a loss in the surface colour appearance of the screen accompanied the form continuity of the figure. If this observation was valid, then any method of preserving the surface colour appearance of the screen should destroy the figural completion. It could easily be tested: (1) by increasing the distance between figure and screen since greater spatial separation should strengthen the segregation of both objects; and (2) by the use of screens of pronounced microstructure. In order to carry out the former experiment (Experiment 11), it was necessary to view

the stimulus situation directly and not on the ground glass of the camera. Suitable measures were taken to avoid shadows on the disc as far as possible. A series of observations, made monocularly (but with good depth impression) by the staff at a distance of 2 m. from the disc, showed that the contours of the figure were still continuous when the depth separation was 1·5 cm. When the distance was increased to 16 cm. and 32 cm., all trace of figural continuity disappeared.

The results of reinforcing the surface colour appearance of the screen by the second method are given in the following experiment.

Experiment 12.

Experimental conditions. A bluish-grey screen with black letters printed on it was fixed to the ground glass. It was placed on the left side of the ground glass and presented simultaneously with the plain standard screen. The visible area of the ground glass was as in Experiment 6.

Subjects and procedure.—The subjects were the 19 practised observers of Experiment 6. The instruction was the same as in that experiment.

Results.

Table V summarizes the findings of this experiment.

TABLE V

A Comparison of the Frequencies of Form Completion Responses with a Structured and Standard Plain Screen

Screen	*N*	*Completion* f	*Completion* %	*Interruption* f	*Interruption* %	? f	? %
Structured	19	8	42	8	42	3	16
Standard plain . .	19	18	95	0	0	1	5

A chi-square analysis of the frequencies of the completion responses with Yates' correction for continuity shows that the lettering on the screen obstructs the form continuity of the figure. Chi-square is equal to 8·1 which is significant at the 0·01 level.

A repetition of the experiment with a small fixation point on the bluish-grey screen brought no further reduction in the number of responses.

The influence of the shape and movement of the figure.

Rosenbach and Metzger accounted for the apparent transparency of the screen largely in terms of the shape, size and movement of the figure. Considerable attention was given to the first of these configural conditions at the beginning of the research. Triangles, hexagons, and circles were tried, but the rectangular form was found to be most satisfactory for our purpose. The remaining conditions, which both Rosenbach and Metzger alleged to be the primary factors in the form completion of the figure, were the object of our final experiment.

The width of the figure.

The relation of the width dimension of the figure to its form continuity is demonstrated in the following experiment.

Experiment 13.

Experimental conditions. The figures were of standard grey but of different widths: 4 mm., 12 mm., 25 mm., 40 mm., and 60 mm. A small dot was placed on the screen as a fixation point. The screen had a reflection factor of 0·19.

Subjects and procedure. The subjects were 5 members of the staff (Mi, Mo, Kn, Ge, Gl). The subjects made four observations on different days. All the figures were shown at each session, but in random order. The subject made his report after two presentations of the figure.

Results.

Table VI gives the results of this experiment.

TABLE VI

FREQUENCIES OF FORM COMPLETION RESPONSES WITH INCREASING WIDTH OF THE FIGURE

Width of the Figure (mm.)	*N*	*Absolute frequency*	*Per cent.*
4	20	3	15
12	20	19	95
25	20	16	80
40	20	9	45
60	20	8	40
	100	55	

If (55) (20)/100 is taken as the expected frequency of continuity responses for all widths, χ^2 is equal to 15·08 which is significant for 4 d.f. beyond the 1 per cent. level. Even with responses on the 4 mm. screen omitted from the Table, the χ^2 value is 12·2 (expected frequencies (52) (20)/100) is significant for 3 d.f. beyond the 1 per cent. level. The low percentage of continuity responses in the case of the 4 mm. screen and those of 40 mm. and 60 mm. resulted from different factors. The small size of the screen's area lying between the projecting parts of the figure at the crossing-point may have been the cause of the large number of negative responses in the first case. With screens of 40 mm. and 60 mm., the subjects could see the upper border and then the lower border of the figure as uninterrupted, but they could not generally see both as continuous simultaneously. It was to avoid confusing such reports of figural continuity with those which implied a simultaneous apprehension of both borders as continuous that a fixation point was placed on the screen. While this avoided one difficulty, it created another: it enhanced the surface colour appearance of the screen. This, as we have already noted, is detrimental to the figural completion.

To this experiment on the width of the figures we may append another concerning the induction effected on the screen when its upper half was presented against a stationary black background, and its lower half against a stationary white background (Experiment 13). The members of the staff unanimously reported a simultaneous brightness contrast effect: the upper portion appeared brighter. The border line separating the two induction fields was then given a swinging movement. Under this new condition, the observers reported a darkening which extended some distance upward on the screen from the point where the border of the figure swept past it. Some saw the darkening on the entire upper half of the screen. The study of the contrast reversal was pursued in a further experiment (Experiment 14). The figure was a dark grey area whose width varied constantly between 3 cm. and 6 cm. Thirteen untrained subjects were asked to observe the area of the screen at the crossing-point and report whether it appeared darker or brighter than the rest of the screen.

Three subjects reported no change whatever on the screen. Of the remainder 4 subjects saw the area at the crossing-point as darker than the other parts of the screen. The other 5 subjects reported the opposite.

These last experiments show that simultaneous brightness contrast is not compelling when the border lines of the inducing fields are in motion. Instead there may be an induction in the opposite sense. It is our opinion that the darkening, which consistently figures in the reports of the subjects under standard conditions, must be related to this effect of an inducing field which is in movement.

The extension of the figure on both sides of the screen.

A series of experiments were undertaken to discover the influence of the length of the figure on its form continuity. These experiments led to the same conclusion. Hence their results can be reported together.

Experiment 15.

Experimental conditions. The figures were standard greys with lines drawn on them in black ink. On one, the lines were vertical and were set 1 cm. apart. On the second, the lines were the same distance from each other but were slanted so that their entire length was 3 cm. They thus exceeded the width of the screen. The linear speed of the figure was reduced to 9 cm. sec.

Subjects and procedure. Four members of the staff (Mi, Kn, Ge, Gl) and two foreign psychologists (B and G), visiting the laboratory, served as subjects. The latter were given a series of practice trials with the plain standard figures before taking part in the experiment. Multiple presentations were given.

Results.

All the observers reported that the lines of 3 cm. length were continuous across the screen, whereas they saw nothing similar in the case of the vertical lines of the second figure. The darkening at the crossing-point appeared completely uniform.

Experiment 16.

Experimental conditions. The figure had a pattern of deep brown stars arranged quincuncially on a light background. Three stars formed a kind of diagonal which exceeded the width of the screen. The other conditions were as in the preceding experiment, except for the screen, which was black.

Subjects and procedure. These were the same as in the preceding experiment.

Results.

The four members of the staff reported that the star pattern could be seen faintly through the screen. The visitors saw some change in the darkening at the critical area which they described as a "flicker" or "irregularity" on that part of the screen.

Experiment 17.

Experimental conditions. A black cross of standard with (the length of the vertical arm was 15 cm. and that of the horizontal arm, 17 cm., the reflection factor was 0·04) was attached to the ground glass as a screen. The remaining area of the ground glass was left uncovered. The figure was a stripe of standard grey brightness (width, 4 mm.), which appeared on the right side of the ground glass as a tilted line. Its tilt gradually diminished as it moved across the field so that at the centre of the ground glass it was in a vertical position and thus covered by the vertical arm of the cross. As it moved across the left side of the ground glass, it again began to tilt, but this time in the opposite direction. The parts of the figure were thus visible on either side of the horizontal bar of the cross except at the middle of the ground glass. The translation speed of the figure was reduced to 4-5ths of its standard speed.

Subjects and procedure. Three members of the staff (Mi, Ge, Gl) served as subjects.

Results.

The reports of all the subjects were identical. The figure was seen as completed across the horizontal bar of the screen when parts of the figure extended on either side of it. As soon as it passed behind the vertical bar of the screen, all trace of the figure disappeared.

Conclusion.

Experiments 15, 16 and 17 fully confirmed that for a figure to be perceptually completed across the screen, its length had to exceed the width of the screen. The necessity of this configural condition appeared most strikingly in the case of Experiment 17 where the subjects could see continuity and interruption successively as the figure approached and passed the vertical bar of the screen. Although the natural expectancy was to see the stripe passing across the whole surface of the ground glass, it completely disappeared in the middle of the field.

The movement of the figure.

The first experiment on this factor consisted in the suppression of the apparent pivotal movement of the figure when it passed the screen (Fig. III). Arcs of 40° of standard grey brightness were used instead of the standard, tangentially placed, rectangular figures. The members of the staff, who served as subjects, reported that the figures were completed across the screen, but added that the continuity was less impressive than when the figure seemed to make a pivotal movement while crossing the screen.

A second experiment was undertaken to study the influence of the speed of the movement of the figure on its form completion.

Experiment 19.

Experimental conditions. We compared the impressions which were given when the figure was stationary with those which arose when it was set in movement at the following linear speeds (cm. sec.): 1·1, 3·5, 9, 18, 28, 47, 53, 68, 76. (All the brightness conditions were standard.)

Subjects and procedure. The subjects were those of Experiment 6. They were instructed to report whether the figure appeared continuous or interrupted in form for each change of speed. The first setting of the figure was at 0. Its speed was then increased successively to 76 cm. sec. A second descending series followed. Thus for all settings there were 38 responses except in the case of 0 and 1·1 cm. sec. Ten subjects reported on the first setting and 18 subjects on the second.

Results.

The frequencies for the various kinds of responses are given in Table VII. If the expected frequencies are computed by taking percentages of the total number of completion responses, the χ^2 value is 11·45. This gives a probability equal to 20 for 8 d.f. In consequence there is no proof that any particular speed significantly favours the continuity response. It should be noted, however, that this conclusion is based on the reports of subjects who were familiar with the experimental task. The distribution of the frequencies suggests that the mean speeds might be significantly more favourable than the others if the experiment were performed in less practised subjects. This hypothesis is supported firstly, by some informal trials with subjects of this kind, and, secondly, by the observations of the members of the staff. The reports show that the form completion was most in evidence with mean speeds of the figure. With faster speeds, they noted that it degenerated into an unshaped darkening.

These experiments on the configurational conditions of the figure and its movement indicate that the projection of the parts of the figure on either side of the screen and their participation in a common movement across the field are the unifying factors principally responsible for the extension of the brightness situation, existing between the figure and the background, to the screen. We have not been able to discover any specific factors, apart from repetition, expectancy and possibly speed which would account for the transition from the initial vague darkening (contrast reversal) to the responses of shaped darkening and of transparency. We have had occasion to notice, however, that the latter response occurs generally with trained observers and is strongly influenced by past experience.

TABLE VII

Frequencies of Form Completion, Interruption, and Doubtful Responses with Increasing Linear Speed of the Figure

Linear speed (*cm. sec.*)	*N*	*Completion*		*Interruption*			
		f	%	*f*	%	*f*	%
0	20	8	40	7	35	5	25
1·1	36	16	44	19	53	1	3
3·5	38	18	47	19	50	1	3
9·0	38	28	74	10	26	0	0
18·0	38	30	79	7	18	1	3
28·0	38	32	84	6	16	0	0
53·0	38	28	74	7	18	3	8
68·0	38	23	60	12	32	3	8
76·0	38	21	55	10	26	7	18
	322	204					

V

Conclusion

We are now in a position to evaluate our initial hypothesis that the "screen effect" and the "tunnel effect" might be related to Rosenbach's case of apparent transparency. If such a relation exists, it is to be expected that the stimulus conditions of all these phenomena would be, to a large extent, similar. Our systematic study of the apparent transparency shows that there are only two such conditions which it possessed in common with the "screen effect" and the "tunnel effect": firstly, the system of stimulation must be such that the common separation line between the figure and the screen appears to belong exclusively to the latter; and, secondly, the visible parts of the figure must lack adequate closure and demand a completion to appear as closed forms. There is an additional condition which the apparent transparency shares with the "tunnel effect." Both phenomena require a narrow screen with movement of the figure on either side of it. Apart from these similarities, the stimulus conditions of the Rosenbach phenomenon are entirely different from those of the two others. These points of difference may be briefly summarized as follows:

(1) The brightness gradient between the figure, screen and background is of primary importance for the apparent transparency while it has practically no influence on the "screen effect" and the "tunnel effect."

(2) A thin line or point can serve as the moving object in "screen effect" and "tunnel effect." This is not true in the Rosenbach phenomenon where the width dimension of the figure contributes to its form completion.

(3) A "screen effect" or "tunnel effect" can readily be had with outlines of figures. On the other hand, it is extremely difficult to see such figures as completed in form under Rosenbach's conditions.

(4) The extension of the figure across the screen has no part in the "screen effect." The impression of "sliding behind" can be given by progressively diminishing the length of the moving figure as soon as it has attained one side of the screen.

(5) The speed of the object's [translation] movement is only of relative importance in the Rosenbach phenomenon, whereas it is a basic condition of the continuity of the movement in the "tunnel effect."

This comparison of the objective factors which control these phenomena, indicates that the relationship between them is more apparent than real. A further consideration will establish this clearly. When naïve subjects are confronted with the stimulus situation proper to the "screen effect" or the "tunnel effect," they spontaneously give an appropriate response. The objective situation is coercive. The apparent transparency, on the contrary, demands an adequate "set" and a certain period of training even under optimum physical conditions. This need of preparation, in addition to the complex stimulus pattern which it requires, is sufficient ground for concluding that it represents a form of structural organization entirely distinct from the others. There is thus no real continuity between the "modal presence" of the hidden part in the Rosenbach phenomenon and its "amodal presence" in the "screen effect" and the "tunnel effect."

This article is an extract from a dissertation submitted to the University of Louvain. The writer is deeply indebted to Professor A. Michotte who directed the investigation.

References

Burke, L. (1952). On the tunnel effect. *Quart. J. exp. Psychol.*, **4**, 121–138.

Knops, L. (1947). Contribution à l'étude de la "naissance" et de la "permanence" phénoménales dans le champ visuel. *Miscellanea Albert Michotte.* Louvain.

Metzger, W. (1936). *Gesetze der Sehens.* Dresden.

Michotte, A. (1950). A propos de la permanence phénoménale: faits et théories. *Acta Psychol.*, **7**, 298–332.

Rosenbach, O. (1902). Zur Lehre von den Urtheiltäuschungen. *Z. Psychol.*, **29**, 434–448.

Sampaio, A. C. (1943). *La Translation des Objets comme Facteur de leur Permanence Phénoménale.* Louvain.

THE QUARTERLY JOURNAL OF
EXPERIMENTAL PSYCHOLOGY

Vol. VI 1954 Part 4

THE NATURE OF SET-TO-LEARN AND OF INTRA-MATERIAL INTERFERENCE IN IMMEDIATE MEMORY

BY

J. BROWN

From the Psychological Laboratory, University of Cambridge.

The influence of set on learning has previously been demonstrated under conditions where set could affect either the selection of stimuli in perception or their subsequent rehearsal. With both these possibilities excluded, set was still found to have a strong influence on learning: this suggests that set can somehow facilitate the formation of the memory trace when a stimulus is perceived.

Two forms of intra-material interference were shown to occur during the recall period. The attempt to recall one part of the material presented interfered with, and was itself impaired by, the attempt to retain the part yet to be recalled.

I

Introduction

A number of experiments have shown that "anticipation" or "set" can have important effects on learning (e.g. Postman and Jenkins, 1948); yet the ways in which these effects may be produced have not been separated. There are several *indirect* ways in which the set of the subject, at the time stimuli are presented, may influence what learning occurs. What is learned must first be perceived, and which stimuli or aspects of stimuli are perceived is largely a function of set. Set may also affect learning through determining which stimuli or aspects of stimuli are rehearsed and organized subsequent to their presentation. But it is also possible that set *directly* affects learning. Set-to-learn may somehow strengthen the memory trace set up at the time of perception. It is this possibility which is studied in the first experiment.

Intra-material interference is sometimes ascribed to pro-active and retro-active inhibition, terms which are only names for the phenomena to be explained. The problem is the *nature* of this interference. In the immediate memory situation, it has a number of possible sources. Kay and Poulton (1951) obtained evidence which suggested that active retention of earlier items is a source of interference with the learning of later items. The present experiments carry the analysis of intra-material interference in immediate memory a step further by demonstrating that recall of items recalled first interferes with the retention of items yet to be recalled; and, conversely, active retention of items yet to be recalled interferes with the recall of earlier items.

II

Method

In the first experiment, two series of items, one of arrows (the A series) and one of numbers (the N series), were presented to the subject in each trial. The subject then attempted to recall either (i) the A series only, or (ii) the N series only, or (iii) first series A and then series N, or (iv) first series N and then series A. He was told which of these alternatives was required either just before or just after the presentation of the series, so that recall requirements were either known (k) or ambiguous (a) during presentation. In all, therefore, eight conditions were used, which can be symbolized A(a), A(k), N(a), N(k), AN(a), AN(k), NA(a), NA(k). Thus A(a) denotes that only series A was to be recalled but that the subject did not know this until after the two series had been presented; NA(k) denotes that series N was to be recalled first and then series A, the subject being told this before the presentation of the series; and so on. Each subject received trials under all conditions: the conditions were given in random order.

The effects of different kinds of set on learning either or both series is shown by the level or recall under each (k) condition as compared with the corresponding (a) condition. For example, comparison of A(k) with A(a) reveals whether set to learn only the A series increases the recall of this series, since only under the former condition was the subject aware during presentation that the N series would not have to be recalled. Perception of both series was ensured under every condition by requiring the subject to make a differential response to every item as it was presented. Rehearsal of earlier items during the presentation of later items was minimized by a fast rate of presentation.

The effect of interpolated recall of N on the recall of A is obtained by comparing the recall of A under conditions A(a) and NA(a). The effect of concurrent retention of N on the recall of A is obtained by comparing conditions A(a) and AN(a). Similarly, the effects of interpolated recall on concurrent retention of A on the recall of N are obtained by comparing conditions N(a) and AN(a) and conditions N(a) and NA(a) respectively. Only (a) conditions can be used for these comparisons, otherwise differential learning during presentation may affect the results.

The second experiment was similar, except that only (a) conditions were used and the two series of items were numbers N and letters L instead of numbers and arrows. Thus the conditions in this experiment can be symbolized N(a), L(a), NL(a) and LN(a).

III

Apparatus

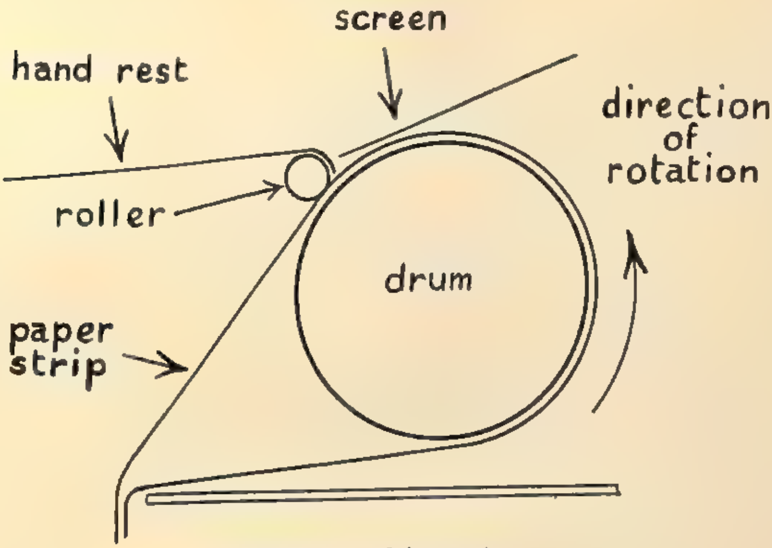
The apparatus may be of some general interest, as it has provided a simple and flexible way of presenting stimuli and of recording hand responses in a variety of experiments. Stimuli are recorded on the left-hand side of a strip of paper and are carried by a rotating horizontal drum behind a screen containing a viewing slit (Figure 1A). The drum is driven by an electric motor controlled by a switch. On the right, the screen has a response slit in which the subject can draw lines to right or left of varying amplitude in response to signals appearing in the viewing window: because the paper is moving, the slope of the lines indicate whether they were drawn in the correct direction (Figure 1B and 1C). If each subject uses a different coloured biro pen or pencil, six or more subjects can be tested using the same paper strip. Scoring is particularly easy, since stimuli and responses appear side by side. The rate of presentation is varied either by altering the spatial distance between successive stimuli or by changing the speed of rotation of the drum. The width of the viewing slit can be altered by a shutter.

In the first experiment, the paper speed was 18 mm./sec. and the interval between successive groups of stimuli 14 mm.: in the second the corresponding values were

9 mm./sec. and 7 mm. Thus in both experiments one group of stimuli was presented every 0·78 sec. The width of the viewing slit was 4·5 mm., just greater than the height of the stimuli.

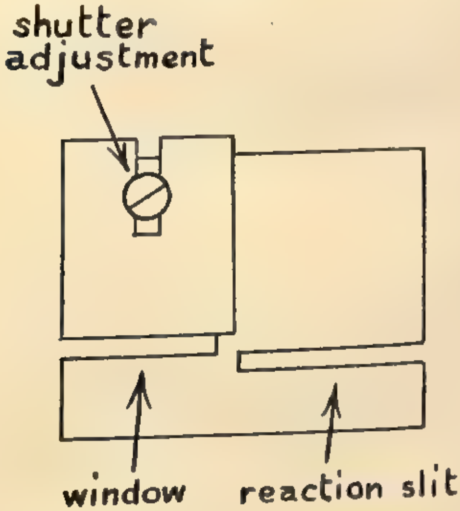
APPARATUS

FIGURE 1A



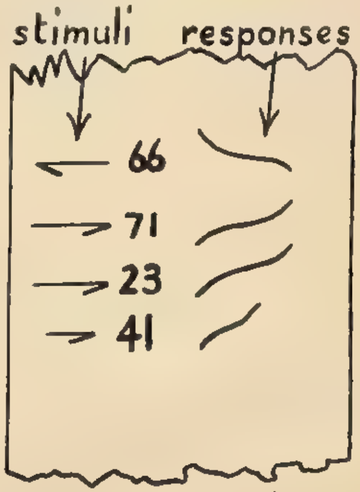
Side View of Apparatus

FIGURE 1B



Screen

FIGURE 1C



Section of paper strip

IV

First Experiment

Details of method.

In each trial, four groups of items were presented to the subject at intervals of 0·78 sec. Each group consisted of a pair of digits and a four-choice arrow; thus two series of items were presented concurrently. Each arrow was long or short and pointed right or left

(Figure 1C). As each group appeared, the subject drew a line in the reaction slit in the direction shown by the arrow and of roughly appropriate length. At the same time, he read out the pair of numbers, which were checked by the experimenter against a key. The stimuli were randomly selected for each trial, except that very easy sequences were eliminated.

Six naval subjects were tested. Each was given four series of eight trials. The eight conditions were given in a different random order in each series and different for each subject, each condition appearing once in each series. The stimuli for each series were recorded on separate paper strips, so that there was a short interval betwen series during which the next strip was inserted into the apparatus. Each strip contained stimuli for one extra trial, for the occasional instances when the subject misunderstood or forgot the recall instructions, thus enabling the condition in question to be unobtrusively repeated.

Under (k) conditions, before switching on the motor, the experimenter told the subject the recall requirements, for example, "Numbers then arrows." Under (a) conditions, this information was given at the moment the motor was switched off at the end of the presentation. The subject recorded his recall attempts on paper.

Subjects were given a practice day one week before the test day.

TABLE I

RECALL SCORES FOR NUMBERS (EXPERIMENT 1)

$N(k) = D_1R_1$	$N(a) = D_2R_1$	$NA(k) = D_1R_2$	$NA(a) = D_2R_2$	$AN(k) = D_1R_3$	$AN(a) = D_2R_3$
150	98	118	70	86	74

(Scores out of 192)

TABLE II

ANALYSIS OF VARIANCE FOR TABLE I

Source	*S.O.S.*	*D.F.*	*M.S.*	*V.R.*	*Significance*
Between S's (subjects)	344	5	68·5	4·87	$P < 0.05$
" R's	273	2	136·5	9·68	< 0.01
" D's	284	1	284·0	20·14	< 0.001
S × R interaction	271	10	27·1	1·92	not significant
S × D "	103	5	20·6	1·46	not significant
R × D "	145	2	72·5	5·14	< 0.02
Residual	141	10	14·1	—	
Total	1561	35			

Comparison	*Difference*	*S.E.*	*"t" (with 10 D.F.)*	*Significance*
N(k) − N(a)	52	13·00	4·000	$P < 0.1$ if $t > 1.812$
NA(k) − NA(a)	48		3·692	$P < 0.05$ if $t > 2.228$
AN(k) − AN(a)	12		0·923	$P < 0.01$ if $t > 3.169$
N(a) − AN(a)	24		1·846	
N(a) − NA(a)	28		2·154	

Results.

Each number or arrow was scored correct only if placed by the subject in the correct position in the sequence. Table I shows the pooled scores for the recall of numbers for the group as a whole. Table II shows a three-way analysis of variance of the results summarized in Table I. The D classification refers to whether the

recall requirements were known during learning, i.e. (k) versus (a) and the R classification to whether the series concerned (in this case numbers) was recalled by itself or before or after the recall of the other series, i.e. N versus NA versus AN. The same trends in the comparisons were found on the practice day. For two of the three comparisons, recall is considerably higher under the (k) than under the (a) condition ($P < 0{\cdot}01$), and in the third comparison the difference, though smaller, is in the same direction.

Comparison of conditions N(a) with conditions AN(a) and NA(a) shows the effects of interpolated recall and concurrent retention respectively. The expected trends are present, but they are significant (at the $P = 0{\cdot}05$ level) only if their significance is measured by a one-tail test (in which case the values shown in Table II of P for different values of "t" should be halved). A one-tail test can be justified on the grounds that the hypothesis of interference predicts the direction of any differences. However, since no similar trends were obtained in the recall of arrows (see Table III), the second experiment was performed in order to obtain more convincing evidence that these forms of interference can play an important role in immediate memory. The significant interaction in Table II ($R \times D$) is that between the recall requirements and whether these were known during presentation.

Table III shows the recall scores for arrows. A three-way analysis of variance revealed no significant differences between conditions. The trends in the (k) versus (a) comparisons are similar to those in the comparable comparisons in the recall of numbers and they were also obtained on the practice day. With arrows, however, the difference between NA(k) and NA(a) is not merely smaller than the other two differences, but is actually negative. Thus (k) seems to favour mainly the series of items to be recalled first.

TABLE III
RECALL SCORES FOR ARROWS (EXPERIMENT 1)

A(R)	A(a)	AN(R)	AN(a)	NA(R)	NA(a)
50	36	49	37	47	38

(Scores out of 96)

V
SECOND EXPERIMENT

Details of method.

The second experiment differed from the first chiefly in employing only (a) conditions. Other differences were as follows. The two series were presented successively instead of concurrently. In each trial, instead of the number-arrow groups, two pairs of number digits followed by a single digit were presented and then two pairs of letters (*consonants*) followed by a single letter. These six groups of stimuli appeared at intervals of 0·78 sec. Each subject received three series of eight trials, each condition occurring twice in each series. (The order of conditions was randomized as in the first experiment.) A fresh group of twelve naval subjects was used. They were given a practice day one week before the test day.

Results.

Table IV shows the recall scores for letters and numbers under each of the conditions for the group as a whole. Table V gives a three-way analysis of the results summarized in Table IV. The R classification concerns the conditions of recall, as in Table II: the B classification refers to whether the recall of letters or numbers

is under consideration. In this experiment, the effects of both interpolated recall (comparison 2, Table V) and concurrent retention (comparison 1) are significant ($P < 0{\cdot}01$ and $P < 0{\cdot}02$ respectively). Comparisons 3–6 show the effects of these interferences when the recall of numbers and of letters are considered separately.

TABLE IV

RECALL SCORES (EXPERIMENT 2)

Recall of numbers ..	N(a) = B_1R_1 176	NL(a) = B_1R_2 150	LN(a) = B_1R_3 142
Recall of letters ..	L(a) = B_2R_1 267	LN(a) = B_2R_2 230	NL(a) = B_2R_3 141

(Scores out of 360)

TABLE V

ANALYSIS OF VARIANCE FOR TABLE IV

Source	*S.O.S.*	*D.F.*	*M.S.*	*V.R.*	*Significance*
Between S's (subjects) ..	458	11	41·6	3·25	$P < 0{\cdot}01$
" R's ..	527	2	263·5	20·58	
" B's ..	393	1	393·0	30·70	
B × R interaction ..	212	2	106·0	8·28	
B × S interaction ..	199	11	18·1	1·41	not significant
R × S interaction ..	233	22	10·6	0·83	
Residual	281	22	12·8	—	
Total	2303	71			

	Comparison	*Difference*	*S.E.*	*"t" (with 22 D.F.)*	*Significance*
1	$R_1 - R_2$	63	24·78	2·542	$P < 0{\cdot}10$ if $t > 1{\cdot}717$
2	$R_1 - R_3$	160		6·457	$P < 0{\cdot}05$ if $t > 2{\cdot}074$
3	N(a) − NL(a)	26	17·53	1·483	$P < 0{\cdot}02$ if $t > 2{\cdot}508$
4	L(a) − LN(a)	37		2·111	$P < 0{\cdot}01$ if $t > 2{\cdot}819$
5	N(a) − LN(a)	34		1·939	
6	L(a) − NL(a)	126		7·188	

It will be seen from Table V that the interaction between B (which recall series) and R (which recall condition) is significant. This is traceable to a smaller effect from the interpolated recall of letters on the recall of numbers than from numbers on the recall of letters.

VI

DISCUSSION

The nature of the effect of set of learning.

Recall tends to be considerably higher under (k) conditions (recall instructions given before presentation) than under corresponding (a) conditions (recall instructions given after presentation). Before this is taken as showing that the "set" of the subject can directly affect learning, the possible influence of other factors has to be considered. The discussion will proceed initially in relation to the largest difference between (k) and (a) conditions, namely, that between N(a) and (N(k): recall was over

50 per cent. higher under the latter condition. Since responses were required to both the arrows and the numbers as they were presented and errors in reponse were very rare, the difference cannot be attributed to the subject ignoring the arrows in the sense of failing to perceive them. Some of the difference may be due to selective rehearsal of the numbers during presentation, since it is impossible to be certain that rehearsal was prevented completely. However, it is unlikely that much of the difference can be so explained. Presentation was at the highest rate at which the subjects could make responses to the items without error: in the language of "Information Theory," rate of gain of information was near the maximum (the rate was actually just under twelve bits per second). The effect of a further factor must also be taken into account. An inevitable feature of the experiment was that recall instructions intervened before recall under (a) conditions but not under (k) conditions. It is therefore probable that this interpolation of instructions before recall contributed to the lower recall under N(a). But it is unlikely to have contributed much since, firstly, recall instructions were brief, viz. "numbers only," and constituted a much smaller interpolation than the recall of arrows before numbers under AN(a); secondly, despite this, the difference in the recall of numbers under N(a) and N(k) is over twice as large as the difference under N(a) and AN(a) (see Table I). It therefore seems reasonable to conclude that other factors can only account for a small part of the difference in recall under N(k) and N(a), most of which can be attributed to a direct effect of set to learn N only on the strength of the traces set up by perception.

The interpretation of other (k) versus (a) comparisons will now be considered briefly. Recall of numbers was significantly higher under NA(k) than NA(a), the difference being almost as large as that between N(k) and N(a): but the recall of arrows was *lower* under NA(k) than under AN(a). Thus there is no evidence here that set to recall one series first resulted in learning being so organized that both series were better recalled. Although differences in the recall of arrows under (k) and (a) conditions were not significant, there was a consistent trend in favour of (k) on both practice and test days, except where numbers were recalled first. It seems probable, therefore, that set can affect the learning of non-verbal material (arrows) as well as of verbal material (numbers).

The results imply that there is some process by which the retention of a limited quantity of material can be facilitated, not depending on the strengthening of learning through rehearsal. It would be plausible to suppose that the retention of any stimulus which has novelty value tends to be facilitated in this way. This would provide an explanation of von Restorff's finding (quoted in Woodworth, 1938) that the single item in a list differing from the rest is better remembered.

Intra-material interference.

The experiments clearly demonstrate two forms of interference during the recall period, namely from the interpolated recall of items recalled first and from the concurrent retention of items yet to be recalled.

It might be suggested that the effect of interpolated recall is due simply to the consequent delay before the subject can attempt to recall the remaining items. However, in a previous experiment (unpublished) no drop in recall was produced if subjects were merely required to delay recall for several seconds. An assumption which would explain both interferences is that retention in immediate memory is an active not a passive process (cf. Kay and Poulton, 1951), since recall during the active retention of later items would produce mutual interference.

A peculiar feature of the first experiment is that there is no sign that these interferences lower the recall of arrows. Thus there seems to be a difference between the

susceptibility of verbal and non-verbal material to their influence. A partial explanation of this difference may be that verbal material is easier to rehearse effectively: for interference with rehearsal would then produce a greater relative effect on the recall of numbers. Both interferences, with verbal materials, were sufficiently large to suggest that they are important factors in limiting the memory span. They will also help to determine the serial position effect. In the recall of a series of items, the early items will be subject mainly to interference from concurrent retention and later items mainly to interference from interpolated recall. According to which interference is stronger, the earlier or the later items will be favoured in recall: the actual serial position effect will also depend on any other forms of intra-material interference present.

My thanks are due to Dr. A. Carpenter for much helpful discussion and to Professor G. C. Drew who read the manuscript. I am also indebted to the Royal Navy for supplying the subjects.

References

Postman, L., and Jenkins, W. O. (1948). An experimental analysis of set in rote learning: the interaction of learning instruction and retention performance. *J. exp. Psychol.*, **38**, 683–9.

Kay, H., and Poulton, E. C. (1951). Anticipation in memorizing. *Brit. J. Psychol.*, **42**, 34–41.

Woodworth, R. S. (1938). *Experimental Psychology*. New York, Chap IX, p. 231.

TACTILE-KINAESTHETIC PERCEPTION OF STRAIGHTNESS IN BLIND AND SIGHTED HUMANS

BY

IAN M. L. HUNTER

(*Psychology Department, University of Edinburgh*)

A comparison is made between the tactile-kinaesthetic perception of straightness in twenty blind and twenty sighted subjects. Both types of subject are found to perceive a "plus curved" edge as straight, but this curve is significantly less for the blind than for the sighted, i.e. the blind subjects' judgments correspond more closely to the objective straight. Furthermore, the blind, both as individuals and as a group, are significantly finer in their judgments and more consistent. These results are thought to be explained in terms of the more highly developed organisation of the blind's tactile-kinaesthetic perception.

I

Introduction

The tactile-kinaesthetic perception of straightness in sighted adults was investigated independently by Rubin (1936) and Blumenfeld (1937). Rubin's eight blindfolded subjects were presented, in the frontal-parallel plane, with a horizontal steel ruler which could either be straight or curved by varying amounts concavely or convexly to the subject. The subject's task was to slide his finger-tips to and fro along the upper edge of this ruler and judge whether it was "plus curved" (i.e. its middle bent away from him), "minus curved" (i.e. its middle bent towards him), or "straight." Each subject was given repeated presentations of the ruler variously curved and the average value calculated of those curves judged "straight." There were two experimental conditions depending on whether or not the upper arm was mobile. In the short-arm condition, only the forearm was moved, the elbow acting as centre; in the long-arm condition, the whole outstretched arm was moved, the shoulder now being centre. The main conclusion was that the edge perceived as straight was, in reality, plus curved. Furthermore, though this plus curve was greater under the short-arm than under the long-arm condition, it was always less curved than the arc traced by the fingertips with the outstretched arm (or forearm) as radius and the shoulder (or elbow) as centre. These findings, as Rubin observed, could be interpreted in two ways. Assuming the complete correspondence of perceived and objective straightness, they demonstrated an illusion. Assuming that the curve judged straight should be the arc described by the moving member, then they demonstrated a constancy effect, or what Thouless (1931) might have called a phenomenal regression to objective straightness. In his study, Blumenfeld used, in addition to the horizontal ruler, a device whereby subjects were presented with three vertical rods in the horizontal frontal-parallel position and were required to place the middle one in a straight line with the two outer ones. His investigation, although more fragmentary than Rubin's, arrived at essentially the same conclusions.

Using his technique of the three vertical rods, Blumenfeld compared the performances of sighted adults with those of four blind adults. Three years later, Crewdson and Zangwill (1940) compared one blind with three sighted adults on a replica of Rubin's horizontal ruler. The results of both studies indicated that (1) the differential threshold between straightness and curvature was lower for the blind than for the sighted; and (2) while, in both types of subject, perceived straightness was given by a plus curve, this curve was greater for the blind than for the sighted. However, because of the small number of subjects employed and the considerable individual differences found, these findings were, at best, merely suggestive. It is

the purpose of the experiment reported below to determine whether these two tentative conclusions are confirmed when larger numbers of blind and sighted are compared.

II

Experiment

Subjects

Taking part were twenty blind and twenty sighted subjects, matched for age and intelligence as nearly as possible. The blind aged from 12 yrs. 6 mos. to 18 yrs. 11 mos. with a mean age of 15 yrs. 8 mos. and a mean I.Q. of 113. The sighted aged from 12 yrs. 6 mos. to 18 yrs. 7 mos. with a mean of 15 yrs. 5 mos. and a mean I.Q. of 111. Although seven of the blind subjects were sensitive to light, not one of them had any object vision. Five were congenitally blind, two had become blind in the first year of life, three in the second year, three in the third, one in the fourth, two in the sixth, one in the seventh, one in the eighth, and two in the tenth year.

Some indication of the histories of the blind subjects is given below, the following five items of information being supplied for each. Sex, whether male (M) or female (F). Age at testing, given in years and months. The present state of vision with regard to which every subject fell into one of only two categories, i.e. completely blind (total) and sensitive to light stimulation (light). The cause and age of onset of blindness. Lastly, the age at which the child's present state of blindness was most probably reached: it is to be stressed that the age given is often an approximate one guessed at in the absence of reliable data.

1.	M.	14·10.	Light.	Keratitis at 1 yr.	1 yr.
2.	M.	12·08.	Total.	Glioma at 2 yrs.	2 yrs. 6 mos.
3.	F.	13·06.	Total.	Glioma at 2 yrs.	4 yrs.
4.	F.	13·11.	Light.	Congenital cataract.	From birth.
5.	F.	13·02.	Total.	Congenital cataract.	From birth.
6.	F.	12·11.	Light.	Congenital familial.	From birth.
7.	F.	15·03.	Total.	Iridocyclitis at 1 yr.	1 yr.
8.	F.	16·10.	Total.	Uveitis at 3 yrs.	3 yrs.
9.	F.	13·10.	Light.	Choroiditis at 7 yrs.	7 yrs.
10.	F.	15·03.	Total.	Keratitis at 2 yrs.	2 yrs.
11.	F.	16·04.	Light.	Burning accident at 6 yrs.	6 yrs.
12.	F.	17·06.	Total.	Choroido-retinitis at 7 yrs.	10 yrs.
13.	F.	17·04.	Total.	Uveitis at 2 yrs.	6 yrs.
14.	F.	18·01.	Total.	Optic atrophy after brain operation at 10 yrs.	10 yrs.
15.	F.	17·05.	Total.	Ophthalmia neonatorum.	From birth.
16.	F.	17·00.	Total.	Glioma at 1 yr.	3 yrs.
17.	F.	18·11.	Total.	Ophthalmia neonatorum.	From birth.
18.	M.	12·06.	Light.	Uveitis at 3 yrs.	3 yrs.
19.	M.	15·02.	Light.	Congenital cataract with deterioration.	8 yrs.
20.	F.	15·08.	Total.	Glioma at 1 yr. 6 mos.	2 yrs.

Apparatus

A flexible steel ruler, 205 mm. long and 0·5 mm. thick, was clamped at both ends and in the middle. The end clamps were fixed to a wooden board so that the ruler was held horizontal with its broad surface vertical and its 0·5 mm. edge uppermost. It was along this edge that the subjects were to move their fingertips. The middle clamp, placed so as not to interfere with the subjects' movements, could be moved, by means of a micrometer screw, either towards the subject (giving a minus curve) or away from him (giving a plus curve). A millimeter scale showed the distance by which the middle of the ruler was displaced from the straight.

Procedure

For every subject, whether blind or sighted and blindfolded, the procedure was identical. He was seated at a table, with his elbow placed in a rest on the table's near edge, and instructed to keep his upper arm stationary throughout the experiment. Only the short-arm condition was used. The apparatus lay before him, in the frontal-parallel plane, in such a position that its middle was directly in front of his elbow, and at such a

distance that he could run his fingertips comfortably along the ruler's length. He was then given a series of trials on which the ruler was variously curved and on which he was required to judge whether it was "straight," or whether the middle was "bent away" or "bent towards." On any one of those trials the ruler might be straight or bent with its middle displaced, positively or negatively, by a distance which was a multiple of 2 mm. These curves were presented in randomized order and each subject had a total of about fifty trials. With this type of procedure (i.e. the psychophysical method of single stimuli), each subject should, ideally, have the same number of trials on each of the same stimuli. But since there were large individual differences in the particular curves consistently judged as bent and since the number of trials which could be given each subject was limited, it was recognized that the ideal procedure would be wasteful in many cases. A compromise was, therefore, adopted in which the range of curves varied from subject to subject as required, but was always such that the two extreme curves were consistently judged as, respectively, "bent away" and "bent towards." This meant that the number of trials involving a particular curve varied from subject to subject, e.g. if he required a range of −4 mm. to +4 mm. he would make ten judgments on each of five curves (−4, −2, 0, +2, and +4), but if he required a range of −8 mm. to +10 mm., he would make five judgments on each of ten curves. Throughout, no restriction was placed on the number of fingers used by the subject nor on the speed or frequency with which they traversed the length of the ruler. Half of the subjects in each group used the left hand and half used the right.

Results

It may be discovered whether the performance of the blind group differed from that of the sighted by comparing the values of all the curves judged "straight" by all the subjects in each group. Table I shows this comparison and enables two statements to be made. First, while perceived straightness corresponded to a plus curve in both groups, this curve was significantly greater for the sighted than for the blind, i.e. the blind's tactile-kinaesthetic perception of straightness corresponded more closely to objective straightness. Second, the judgments of the sighted were significantly more variable than those of the blind.

TABLE I

	No. "straight" judgments	*Mean*	*Standard deviation*
Blind Group ..	305	+1·44	3·03
Sighted Group ..	407	+2·37	4·93
t ..		3·10	9·05
p ..		<·01	<·001

The second statement must now be elaborated since the greater variability of the sighted could have been due to three different conditions. (1) They could have had a higher differential threshold between straight and curved. (2) They could have differed more among themselves regarding the mean values of their "straight" judgments. (3) As individuals, they could have been less consistent in their judgments on the same curve. Detecting the presence of the first condition involves some measure of the scatter of the individual's judgments. Such a measure is the "range," defined as the number of curves intervening between those which were consistently judged as either "bent away" or "bent towards." The mean range of the blind subjects is found to have been 5·6 curves with a standard deviation of 1·07; the sighted had a mean range of 7·2 curves and a standard deviation of 2·36. The difference between the two means and the difference between the two standard deviations are both significant at the ·01 level of probability. Whether the second

condition was present can be determined by calculating the mean value of the curves judged "straight" by each individual subject. When this is done the standard deviations of these mean values is 1·15 for the blind and 2·71 for the sighted. The difference between the two is significant at the ·005 level. The presence of the third condition may be detected by calculating, for each group, the proportion of curves on which all three judgments were passed by the same individual. For the blind, this proportion was ·046 and, for the sighted, was ·201. These proportions are significantly different beyond the ·001 level. To conclude: as contrasted with the blind subject, the sighted subject was less consistent in his judgments, had a wider scatter of judgments, and differed more from his fellows with regard to the extent of this scatter and its mean value. In other words, the blind, both as individuals and as a group, were finer in their judgments and more consistent.

The present experiment yields evidence not only on the effect of visual experience (or the lack of it) on the tactile-kinaesthetic perception of straightness, but also on the effect of chronological age. The subjects, blind and sighted together, may be split into three age groups, i.e. fifteen aged 12 yrs. 6 mos. to 13 yrs. 11 mos. (Age Group I), thirteen aged 14 yrs. 0 mos. (Age group II) to 16 yrs. 11 mos., and twelve aged 17 yrs. 0 mos. to 18 yrs. 11 mos. (Age Group III). The performances of Age Group I and Age Group III (the mean ages of which are 13 yrs. 1 mos. and 17 yrs. 10 mos., respectively) are compared in Table II. This table shows that perception

TABLE II

	No. "straight" judgments	*Mean*	*Standard deviation*
Age Group I ..	309	+2·01	4·96
Age Group III ..	243	+1·79	3·83
t ..		·59	4·35
p ..		$>$·5	$<$·001

was no more objective in the one group than in the other, but that judgments were significantly less variable in the older group than in the younger.

No consistent difference was found between those subjects who used their left hand and those who used their right.

III

Discussion

In interpreting the above results, there can be little doubt that a practice effect is responsible for the smaller variability of judgment shown by the blind as opposed to the sighted and by the older subjects as opposed to the younger. The postulation of such an effect rests on three reasonable suppositions: that perceptual judgments of any sort become less variable with practice; that practice in tactile-kinaesthetic perceiving increases with age; and that, because of the greater part it plays in their reading, writing and perception of the world generally, the blind have "more" experience of tactile-kinaesthetic judgment than the sighted. The third supposition might be better expressed by saying that, because they are obliged to depend more exclusively upon it, the blind have a more highly organised tactile-kinaesthetic perceptual system in the same way as the taste of the tea blender and the hearing of the musician is more highly developed than is ordinarily so.

The greater objectivity of the blind's straightness judgments is open to more complex interpretation. At least logically, two possibilities exist. Perhaps the more highly developed the visual perception, the less objective is the judgment of tactile-kinaesthetic straightness. Or it could be that this judgment becomes more objective with the greater development of tactile-kinaesthetic perception. The former possibility seems the less plausible. Rubin has shown, admittedly, that vision may influence tactile-kinaesthetic perception. When a subject explores an edge with his fingers and simultaneously looks at it, the plus curve is no longer judged straight but tactile-kinaesthetic judgment is brought into line with the dominant visual judgment; the edge which "looks" straight also "feels" straight. Furthermore, Rubin reports that, where a blindfolded subject has a knowledge of the objective situation and employs strong visual imagery, his judgments may be the same as those given while vision is actually present. But where a sighted subject is blindfolded it is difficult to imagine just how his superior visual organisation could operate to make subjective straightness correspond more closely to the arc traversed by the fingertips with the elbow as centre. If anything, it ought to make judgment more objective. The second possibility mentioned above seems to fit the facts more readily. However, the growth of tactile-kinaesthetic perception cannot be accepted as the explanation until the lack of difference between Age Groups I and III has been satisfactorily accounted for.

With both blind and sighted subjects, practice between the ages of thirteen and eighteen years has little effect on the objectivity of straightness judgments. This could be explained if the effects of tactile-kinaesthetic practice were restricted to a certain developmental stage or, alternatively, were limited in amount. This would mean, in the one case, that whatever effect practice has on objectivity is, unlike its effect on variability, complete before the age of thirteen. Otherwise expressed, practice is limited to a critical period of development so that, by thirteen years, judgments of straightness, and perhaps other aspects of space perception besides, are "set" to be influenced thereafter only by complex processes of instruction and reasoning. Were this so, persons blinded after the critical period would never give straightness judgments comparable in objectivity to those of persons blinded before the critical period. Unless, that is, they came to understand that the moving arm described an arc rather than a straight line and, by reasoning, compensated for this. The alternative explanation is in terms of an absolute limit to the amount of practice which is effective. The blind might not improve after thirteen because they had already reached the limits of practice. The sighted, on the other hand, might have reached an artificial limit of tactile-kinaesthetic practice temporarily imposed by their use of vision. Were this so, persons blinded at any age whatever would eventually become more objective in their straightness judgments. Which, if either, of those two explanations holds is a point which must be settled on the basis of evidence not yet available. But, as regards the present discussion, the important thing is the demonstration that a reasonable explanation does exist for the results shown by the two age groups.

Perhaps, then, the most plausible explanation of the blind's greater objectivity in the tactile-kinaesthetic perception of straightness is that, as a result of greater practice, they approximate more closely to an objective judgment and, although never completely free from the distorting effect of the arm's radial movement, are less influenced by it. In harmony with this explanation is the finding of Crewdson and Zangwill that their blind subject's perception of straightness was less affected than was that of their sighted subjects by a change from the short-arm to the long-arm condition. It is likely that, had the blind subjects in this experiment all been

blind from birth, the differences between them and the sighted would have been even greater.

This experiment was conducted adjunctively to a more extensive study by Professor Drever. It was he who arranged for subjects to come to the laboratory, and I am sincerely indebted to him for letting me examine them during their visit. To be thanked for making subjects available are the headmasters of Craigmillar Park Blind School, Moray House School, and George Heriot's School. Dr. Semeonoff contributed valuable advice on the use of statistics.

References

Blumenfeld, W. (1937). The relationship between the optical and haptic construction of space. *Acta Psychol.*, **2**, 125–74.

Crewdson, J., and Zangwill, O. L. (1940). A note on tactual perception in a blind subject. *Brit. J. Psychol.*, **30**, 224–29.

Rubin, E. (1936). Haptische Untersuchungen. *Acta Psychol.*, **1**, 285–380.

Thouless, R. H. (1931). Phenomenal regression to the "real" object. *Brit. J. Psychol.*, **21**, 339–59.

THE EFFECTS OF POSITION IN A DISPLAY UPON PROBLEM SOLVING

BY

HARRY KAY

From the Institute of Experimental Psychology, University of Oxford.

Sixty-four subjects, aged between 15 and 72 years were told how to carry out a problem-solving task, whose difficulty could be varied by changing the position of an index card in relation to a display and control.

The outstanding feature, particularly of the older group's performance was the repetition of the same two kinds of mistake. Their failure seemed mainly due to an inability to rid themselves of their wrong "solutions," in spite of the constant information indicating the mistake at each position. Whilst these differences between the age groups were not pronounced at the easiest task, they were at the most difficult.

It is suggested that: Firstly, the procedure of giving a subject the solution to a problem and observing how he reconciles his information with it, though only operating in a circumscribed area of problem-solving, is a genuine example of thinking, and, particularly where certain kinds of mistake can be predicted, it enables the experimenter to gain a more exact impression of what a subject is doing. Secondly, the manipulation of spatial contiguities lends itself to flexible and fairly precise measurement, and the principle might usefully be extended to other variables such as size, time intervals, colour, etc.

I

Introduction

This experiment does not fit neatly into any rigid category of learning or problem-solving. Several of the issues which are raised by this study will only be briefly indicated, and this paper will primarily attempt to give a full factual account of the proceedings, as it is felt that they could serve as a model for further experiments.

A previous study upon serial learning with human subjects (Kay, 1951) had illustrated how a task which would traditionally be defined as rote learning was anything but a mere rote recapitulation of the presented information. It was a feature of adult learning to repeat mistakes—that is, for subjects to ignore information which was continuously being given to them and to adhere to that version of an operation which they had themselves created. This repetitive tendency, expressed as an inability of an organism to unlearn what it has already learned, is a common enough phenomenon of behaviour, and this experiment sought to examine it in a task where as little emphasis as possible was being placed upon remembering.

The aims were to devise an experiment where—

(i) Subjects of different ages had to carry out a particular mental task.

(ii) The probable kinds of error could be largely predicted.

(iii) They would occur sufficiently often to provide quantitative data, so as to give some indication of the mental processes which had taken place.

Because of a previous experiment it was desirable that—

(iv) The task should impose as little stress as possible upon memory.

II

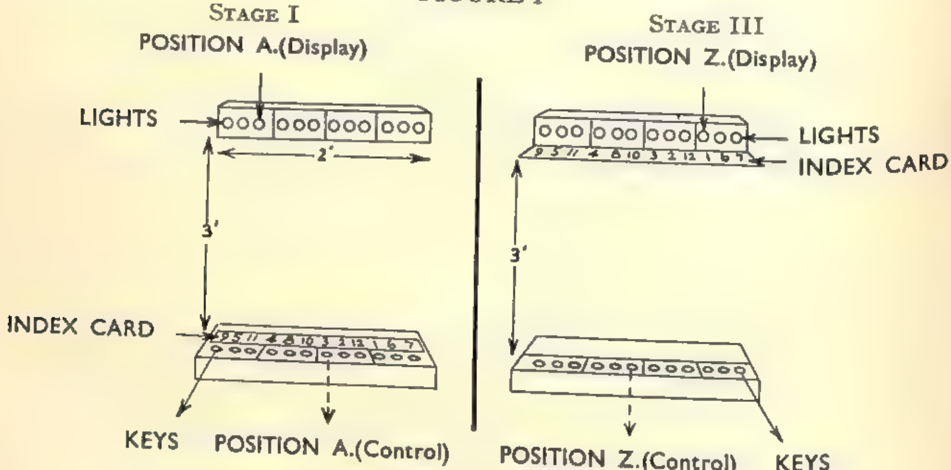
Methods and Procedure

Apparatus

The layout of the apparatus has been sketched in Figure 1. Subjects sat in front of a row of 12 Morse keys, spaced 2 inches apart. A row of lights, spaced in exactly the same way, was placed 3 feet from the keys. An index card with a randomized order of numbers upon it was positioned immediately behind the keys (*see* Figure 1, Stage I). The numerals

on the card were also spaced 2 inches apart so that they were directly in line with both the keys and lights. Markers were put after every third light and key to facilitate the numbering of spatial positions. The apparatus was constructed so that when a "correct" key was pressed it switched out the illuminated light and switched on another. When an "error" key was pressed nothing happened (so far as the subject was aware). There was then immediate knowledge of results after each response. The apparatus was wired up to a 14-pen polygraph recorder so that each key was scored separately, in addition to time intervals and correct responses. The apparatus has been described fully elsewhere (*see* Welford, 1952).

FIGURE 1



Positions of the Light-Key Apparatus at Stages I and III.

Subjects.

Sixty-four subjects took part in the experiments, 10 belonging to each of the age groups, 15–24, 25–34, 35–44, 45–54, 65–72, and 14 to the group 55–64. All subjects in the groups 15–64 were employed in one of two departments in the York factory of Messrs. Rowntree & Co. The group aged 65–72 were pensioners from the factory, who regularly attended the Pensioners' Club on the factory premises. All subjects were volunteers and all were most co-operative and helpful.

Instructions to Subjects.

A subject was seated with the apparatus in front of him and with the index card immediately behind the keys (as in Figure 1, Stage I). After he had been shown how the apparatus worked, he was given the following typewritten instructions:—

1. Think of the lights as being numbered 1 to 12 from the left.
2. When a light comes on decide which number it is.
3. Find that number on the card.
4. The correct key to hit is the key in line with the number selected on the card.

During the reading of the instructions the experimenter occupied himself with the apparatus and tried not to hurry nor to embarrass subjects. When they indicated they had finished studying the instructions the experimenter checked with all subjects if they thought they understood them and then by demonstration and questions—ostensibly to help subjects—he ascertained whether in fact they did understand correctly. These results were scored. The first stage of the experiment began after subjects had not only satisfied themselves that they thought they knew what to do, but had also satisfied the experimenter that they could begin the task correctly.

The same typewritten instructions were retained by a subject and used throughout all stages of the experiment.

Experimental Design.

Subjects were carrying out what was essentially the same task at three different stages, using the same series of 10 lights at each. The series of 10 was given twice at each stage,

so that *each* was measuring the errors and the time taken to make 20 correct responses. (Of the 12 light positions only 10 were used: positions 2 and 11 were omitted from the series, but were used for demonstrating and checking.)

Stage I. The card was placed so that each of its numerals was immediately behind a key (*see* Figure 1, Stage I). After subjects had shown that they knew how to carry out the instructions they began this stage of the experiment.

Stage II. The card was now moved to a position mid-way between the keys and the lights. Subjects were again asked to do the task exactly as indicated on their typewritten instructions, and since all of them had done the first stage satisfactorily it was pointed out to them that this was really asking them to repeat the same procedure as for Stage I.

Stage III. The card was now moved to a position in front of the lights, so that each light had a numeral immediately in front of it. Subjects were asked to continue to carry out the same instructions.

Further stages in the experiment were carried out, but these results are not being discussed in this paper.

It should be mentioned that, though we are speaking of a series of 10 or 20 responses, each response was independent of another. However, since the same Index Card was used for all three stages it was to be expected that some incidental learning of the series might occur. One of the further stages of the experiment (not discussed here) did in fact measure this incidental learning by using a different Index Card. The amount of such learning was relatively small and was less for the older than for the younger subjects. In so far as any learning of the series would tend to lessen the difficulty of Stage III it should have decreased the discrepancies between the results of the three stages which will be discussed.

Experimental Measurements.

At each stage of the experiment there were 20 "light" positions and a subject had to make a correct response at each. Two measurements were recorded:—

1. The time in seconds.
2. The number of mistakes. These have been subdivided into:
   (*a*) *First Errors.* This limits the mistakes to the first error only at each serial position; therefore, a subject can only make 20 first errors at any one stage.
   (*b*) *Total Errors.* Every mistake made at any serial position.

The distinction aids in studying the repetition of mistakes; elsewhere the terms "errors" or "mistakes" will be used to refer to either kind of wrong response.

Explanation of Task.

From the experimental instructions which have just been given it will be observed that in order to carry them out four points had to be determined:—

A. The position of the light.
B. The number of that spatial position of the light from the left. (*Spatial position number.*)
C. The numeral on the index card corresponding to that spatial position number. (*Card numeral.*)
D. The key in line with the card numeral. (*Key position.*)

To complete each response a subject had to carry out a series of substitutions, somewhat similar to that in the Digit Symbol Test of the Wechsler-Bellevue Intelligence Scale. But whereas the Wechsler Substitutions are made with a number of symbols which are learned as quickly as possible, here the aim was for all moves to be governed by some common and simple principles. Once these were learned a subject had nothing further to retain, and the experimenter's problem was to complicate the task without changing its principles. This was achieved by varying the spatial position of the index card in relation to the display and control panels. We may illustrate this by two examples.

Let us say that in Figure 1, Stage I, the light is on at position A. This was the third light position from the left end of the display. Subjects therefore found numeral three on the card and pressed the key in line with that number. (Position A Control.)

It was known that this was a simple procedure; it was easy enough to give the light a spatial position number when there was no index card near and, at the last step, to chose the key nearest to the card numeral when the two were adjacent.

But for Stage III, though the procedure was the same, both these steps were more complicated. The index card was adjacent to the lights so that there was a tendency to associate, incorrectly, a light with its nearest numeral, whilst the numerals were some distance from the keys to which they did refer. The difficulty can be observed in the example for Figure 1, Stage III, where if the four points were correctly determined we have—

A. The light is at Position Z.
B. *Spatial Position Number* of Z is 10.
C. *Card Numerals*, 10 are found.
D. *Key Position* is in line with numerals 10 (Position Z Control).

Particular Categories of Mistake.

The changing of the position of the index card in relation to the display and control made it possible to complicate or facilitate the operation as desired. But it followed that if the difficulty was being caused by a change in spatial contiguities that certain types of mistake would predominate. If their occurrence could be measured it would allow us to be much more precise about the difficulties that had been created. This section, therefore, outlines the two kinds of mistake which were predicted to occur when the index card was moved from the control to the display. (The reader who can predict the mistakes for himself is advised to continue with the results.)

We observed (above) that to carry out the task four points had to be established: this involved a subject in making three substitutions (or transformations) of the following kind:—

(i) The *light position* (A) is transformed into a *Spatial Position Number* (B).
(ii) The *Spatial Position Number* (B) is transformed into the corresponding *Card Numeral* (C).
(iii) The *Card Numeral* (C) is aligned with the *Key Position* (D).

The predicted categories of mistakes will be discussed by referring to this procedure and using the example given for Figure 1, Stage III.

Errors Type I.

It was anticipated that these errors would be caused by a subject assuming that the card numeral immediately in front of the light indicated the spatial position number of the key. The procedure would be—

A. The light is at position Z.
C. The numeral on the card *aligned* immediately below the light is identified as the correct Card Numeral (i.e. No. 1).
D. The key whose spatial position number is that of the Card Numeral is selected. (Key No. 1.)

There are two predominant features of this mistake.

(i) The first transformation is omitted (Spatial Position Number of the light).
(ii) This may arise because the order of the transformations is changed—the procedure begins where it should end with an alignment; that is, alignment with the Key becomes alignment with the numeral on the card.

Errors Type II.

It was expected that a different type of mistake would arise through subjects alinging the light position with the card numeral immediately in front of it, as in the Type I Errors, and then striking the key in line with the numeral.

In the example of Figure 1, Stage III, this mistake would produce the following:

A. The light is at position Z.

C. The card numeral immediately in line with the light is selected (No. 1).

D. The key in line with the card numeral is selected.

This mistake carries out an alignment procedure twice, which has the effect of nullifying the use of the index card. It seems a fairly obvious mistake and it was not anticipated to be as prevalent as that of Type 1. (The experimental results were to indicate the reverse.)

A predilection for either type of error at Stage III would indicate the strength of the spatial contiguity influence and the kind of mental operations which had taken place.

Random Errors.

This category was intended to include all errors not falling into the other two, and carried no further implication than this. It was thought that it would be made up of mistakes caused by misalignments, by wrong numbering, or the occasions when a subject became so "lost" that he hit all the keys in turn until he came to the right one.

III

Results

1. *The Intepretation of Written Instructions.*

Before turning to the main body of results there is an interesting note upon how adequately subjects understood their instructions. It was essential that subjects should not be mistaken about what they had to do before beginning the experiment. Accordingly, as already mentioned, the experimenter checked with subjects after they had finished studying the instructions if they thought they could carry them out, and whether they could in fact do so. Table I gives the results.

TABLE I

Ability of Subjects to Carry Out Written Instructions

Age Group	*Percentage number of subjects in each age group who:*		
	Could carry out the instructions correctly (*i*)	*Thought they could carry them out and had interpreted them wrongly* (*ii*)	*Who could not understand what they had to do* (*iii*)
15–24 ..	60	40	0
25–34 ..	50	50	0
35–44 ..	30	50	20
45–54 ..	30	40	30
55–64 ..	28	44	28
65–72 ..	20	40	40
For all Groups ..	36·3	44·0	19·7

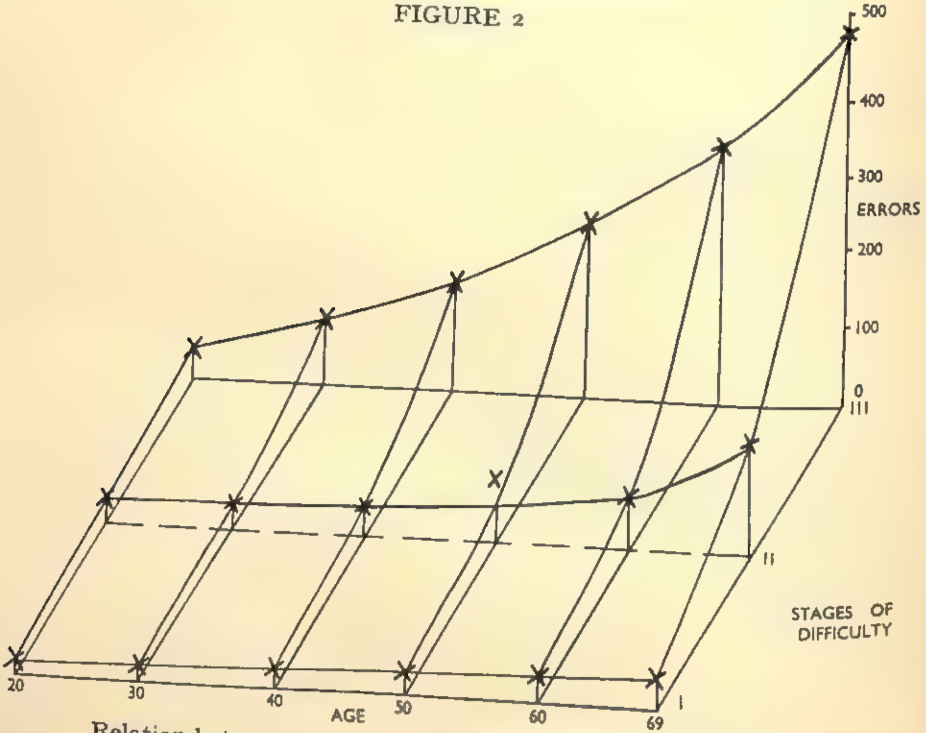
Though the instructions contained only four short simple sentences 44 per cent. of the subjects were prepared to begin the experiment saying they knew what to do when that was just what they did not know.

A second feature of Table I is that the percentage number of subjects who were able to understand the instructions decreases with age. Conversely, the number of those who knew that they did not understand them rises steadily with age (column 3).

2. *Carrying Out a Task at Three Stages of Difficulty.*

The recorded times and mistakes in performing the three stages of the experiment are set out in Table II and the errors are graphed in Figure 2.

FIGURE 2



Relation between Age, Total Errors, and Degree of Difficulty.

Subjects apparently found Stage I a fairly simple operation. They did not take long to carry it out and the number of mistakes made by all the different age groups was small. Even so the two oldest groups were making two and a half times the number of errors and taking one and a half times as long as the two youngest age groups.

At Stage II, results for all ages indicate the greater difficulty of this task over Stage I. It will be remembered that as the index card was now exactly mid-way between the display and the controls, there was nothing in the spatial layout to suggest that it should be associated with one more than the other. In so far as a subject had already followed the correct procedure at Stage I it was to be expected that he would not become confused, but in so far as the spatial contiguities no longer aided the underlying principles as they did at Stage I, it was thought that subjects would experience more difficulty.

In fact all age groups did make one and a half times to twice as many errors as they had done at Stage I, and the time scores showed a similar though slightly more marked trend. (*See* Table II.)

TABLE II

First Errors and Times at Stages I, II and III

Age Group	*Stage I*		*Stage II*		*Stage III*	
	Errors					
	Means	*S.D.*	*Means*	*S.D.*	*Means*	*S.D.*
15–25 ..	1·1	0·94	2·0	1·79	2·7	2·05
25–34 ..	1·3	0·90	2·5	2·06	4·4	1·73
35–44 ..	1·5	0·92	3·0	1·48	5·6	3·22
45–54 ..	2·1	1·30	3·7	2·96	7·6	4·17
55–64 ..	2·7	2·00	4·3	2·76	8·2	5·96
65–72 ..	2·6	1·88	4·4	3·83	12·8	5·27
	Times					
15–25 ..	56·4	9·89	70·8	17·18	84·8	35·21
25–34	54·2	4·79	71·7	18·28	111·6	43·20
35–44	62·0	29·34	85·7	31·97	137·1	84·85
45–54 ..	64·1	14·53	96·3	34·79	174·7	81·69
55–64 ..	73·7	14·39	124·8	67·44	229·3	135·61
65–72	84·7	25·96	198·7	175·90	445·3	226·71

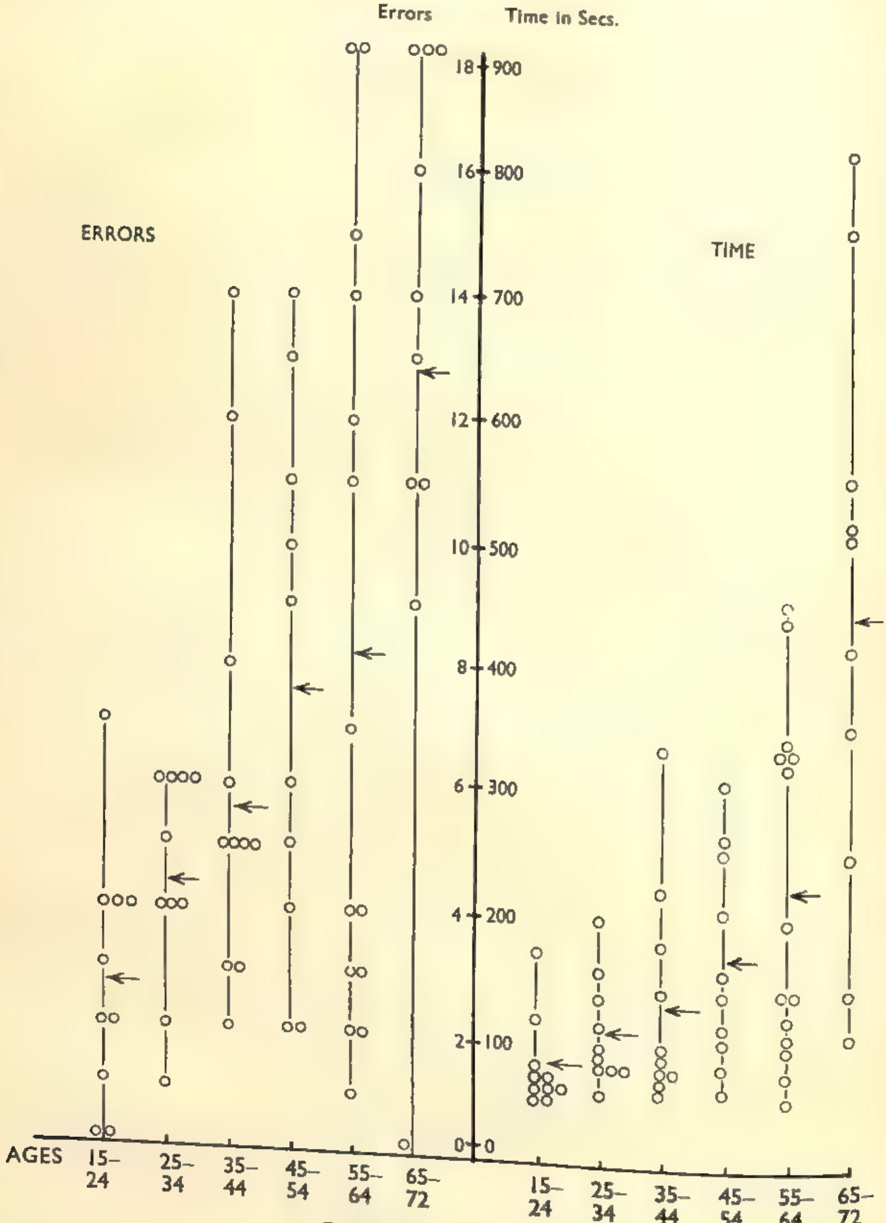
At Stage III, where the position of the index card was immediately in front of the display, the results again indicated the increased difficulty. The youngest groups were now making about three times the number of mistakes and taking approximately twice as long as at Stage I, whilst the older groups were more adversely affected. When the over 65's are compared at this stage with the under 25's they were making more than four times the errors and taking more than five times as long.

If the results at these three stages are considered as indicating the trend which takes place between different age groups as they learn to carry out a task of increasing complexity, there is a definite tendency for the differences between the performances of each age group to increase as the difficulty of the task is increased. (*See* Figure 2.)

Individual Differences.

It will be appreciated that as yet we have spoken only of the result in terms of the mean values for each task, but the standard deviations show that the individual results are more complicated than has so far been made out. As would be expected when the

FIGURE 3



Individual Results at Stage III.

task became more difficult the mean values not only increased, but the variations between individuals became accentuated. The effect of this was particularly marked with subjects in the two oldest groups. In order that as comprehensive a picture

as possible might be given, Figure 3 sets out the individual distributions for Stage III. These results not only indicate the general trend as given by the means, but how many individuals in each group fail to conform to it. This variability amongst older subjects is such a general finding (*see* Thorndike, 1928; Welford, 1951) that it will not be discussed here. It has to be accepted that in learning experiments age is a fickle variable, unequally influencing performance from one individual to the next, but inexorably depressing the standard of a whole group of subjects.

Categories of Mistake.

So far the customary measurements of mistakes and time have suggested that the changes in the position of the index card have considerably increased the difficulty of the task. Now by examining the kinds of mistakes which were made it is hoped to be more precise about the effects. Accordingly, the two types of mistakes already discussed have been analysed separately and are given in Table III.

TABLE III

DISTRIBUTION OF TOTAL ERRORS AT STAGES I, II AND III

Age Group	*Total Errors*	*Errors Type I*	*Errors Type II*	*Random Errors*
		Stage I		
15–24 ..	12	6	–	6
25–34 ..	13	8	3	2
35–44 ..	26	5	5	16
45–54 ..	26	3	4	19
55–64 ..	36	2	4	30
65–72 ..	31	4	5	22
TOTALS ..	144	28	21	95
		Stage II		
15–24 ..	29	2	4	23
25–34 ..	32	7	7	18
35–44 ..	45	8	7	30
45–54 ..	85	17	22	46
55–64 ..	73	17	21	35
65–72	150	50	52	48
TOTALS ..	414	101	113	200
		Stage III		
15–24 ..	40	10	13	17
25–34 ..	85	18	29	38
35–44	136	44	50	42
45–54 ..	235	76	85	74
55–64 ..	336	116	114	106
65–72 ..	479	156	226	97
TOTALS ..	1,311	420	517	374

Errors Type I.

It will be observed from Table III that at Stage I there were not many errors of this first type; 64 subjects only contributed 28 mistakes. But at Stage II there were

101 errors, whilst at Stage III this had risen to 420, a fifteen-fold increase over Stage I. It will be appreciated that the only reason why subjects changed the order of their mental procedures and made this mistake was that the position of the index card was moved. Their "reasoning" as expressed by one subject was, "the lights and the number on the card are close together; when a light comes on it seems to mean the number immediately below it."

Subjects in aligning a light with the card numeral in front of it, were, of course, making an association which was opposed both to the instructions for the task and to how they had hitherto carried them out. This seems to indicate certain priorities in their treatment of data, for having followed out the above procedure they realised that they had not fulfilled their instructions in that a different substitution was uncompleted. So they then, instead of aligning numeral with key, found the key whose spatial position number was corresponding with the numeral on the card, that is they reverted to a version of the substitution procedure which they should have done earlier.

Errors Type II.

It was not anticipated that this type of error would be as prevalent as it in fact proved to be. Table III shows that it occurs even more often than that of Type I. This is surprising since the mistake seems almost too naïvely wrong to have happened so often. However much subjects might be influenced by a tendency to align the lights and the card numerals, it had been argued that subjects would surely be aware that they had omitted some necessary steps in achieving this patent error. However, from watching subjects perform, it was noticeable that these errors occurred more spontaneously than those of the first type—it was often a subject's quick reaction to the problem. As a result the procedure often seemed to be tried, but it did not appear to produce the same degree of bewilderment when it proved to be wrong. Subjects were prepared to accept it as a mistake with remarks such as "Of course that is wrong!" but this did not prevent them offering it as a solution at the next serial position when they were in difficulties. Hence the huge number of mistakes of this type.

Random Errors.

From Table III, it will be seen that whereas at Stage I two-thirds of the errors fall into this category, the figure is less than one-third at Stage III. It is not, of course, being claimed that the analysis of the two major types of mistake covers all the wrong mental decisions which subjects took at the last stage, but it is felt that the random errors are sufficiently small to be accounted for by inevitable minor aberrations, and those occasions when a subject did become so confused that he hit all the keys in succession to try to find the correct one. To substantiate this point it should, perhaps, be mentioned that for any one response the probability of hitting by chance the right key, or a particular one of the two types of errors was 1 in 12, whilst that of making random errors was 9 out of 12. Thus, whenever a subject indulged in making a random selection he was likely to make considerably more random errors than any other kind.

Repetition of Mistakes.

The number of errors at Stage III, given in Table III, would suggest that the same mistake was frequently made by the same subject. This was particularly true of the oldest subjects who must have been repeating over and over again the same two types of errors. To achieve as they did at Stage III 156 "Type I" errors and 226 "Type II" implies that on the average each subject in finding the required 20 correct responses, made nearly 40 incorrect responses of the same two types.

But even so a subject might have repeated most of these two types of errors at the same rather than at different serial positions. The mean number of first errors at this stage, as given in Table II, makes this unlikely for the older groups, but the point can perhaps be settled most clearly by considering an individual record such as that in Table IV. This result is in many ways typical of the subjects in the over 65 age group. It records the first 10 serial positions at Stage III, where a subject made 25 errors, of which 10 were Type I, 14 Type II, and one was random. From the table and its accompanying explanation of the responses, the predominant feature of such a performance was that a subject, after making certain types of errors at one serial position and subsequently amending them by following the right procedure, was unable to continue correctly. Out of the 10 serial positions, at 7 both types of mistake were made, and it will be observed that the three correct positions did not occur consecutively, but were interspaced between the mistakes. It was this ephemeral quality of the solution which was so evident; far from being permanent it was typical of this performance, as of so many, to make a mistake, to be able to correct and then to make just the same type of mistake at the next serial position.

TABLE IV

A Typical Individual Record (Subject aged 68) for Stage III: First 10 Responses

Position of lights in serial order	1	9	3	12	10	6	8	5	7	4
Correct keys in serial order	10	1	7	9	6	11	5	2	12	4
Subject's Responses ..	1* 9† 1* 1* X	I	11† 11† 3* 3* VII	7† 12* 12* IX	10* 1† VI	XI	2† 8* V	5* 8† 8† 5* 8† 5* II	7* 3† 5r 7* XII	IV
Total Errors ..	4	0	4	3	2	0	2	6	4	0

† = Type I errors. (Card numeral equals spatial position number of key.)
* = Type II errors. (Alignment procedures only.)
r = Random errors.
Roman Numerals = Correct keys.

The table should be read in conjunction with Figure I, Stage III.

The subject is presented with Light I. He aligns it with the card numeral and then aligns the numeral with the key position and strikes key 1 (Type II error). He then reads the number on the card immediately underneath the light and strikes the key whose spatial position corresponds to this number (Key 9) (Type I error). He then reverts to his first procedure of taking the light position as being equivalent to the key position and twice hits Key 1. Finally, he goes through the correct procedure of finding the position number of the light, finding that digit on the card and then hitting the key in line with that numeral (Key 10). This changes the light from 1 to 9. Subject here presses the right key, and the light changes to position 3, and so on.

Note how often the subject makes the same two types of mistakes at a serial position, then gets it right, only to repeat the same mistakes at the next serial position.

Summary of Older Subjects' Performances.

Bearing in mind the analysis of the errors of an individual subject which has just been made and the error figures presented in Table III, the total performance of the

over 65 group can be summarised as follows. Each of these 10 subjects had to make a correct response at 20 positions of the lights, making a total of 200. Of these 200 responses 128 were wrong. (First Errors, Table II.) Thus, as this mean for the errors is fairly representative, it can be said that a subject in this group in making 20 correct moves made mistakes at nearly 13 different positions, and repeated at least one of the same two types of mistakes at each of these 13.

The results of the oldest age group have been given in some detail because they do represent an extreme form of a tendency which is evident enough amongst all the older groups. Faced with the problem of how to carry out their instructions, subjects were putting forward consistently two types of responses which were proving wrong; and, as with the oldest age group the outstanding characteristics was that in spite of always having to find, and in fact finding, the correct response before moving to another serial position, they were subsequently repeating these same two types of mistake. Had subjects been able to accept the evidence, both positive and negative with which they were confronted, then they should have been able to manipulate the variables involved in the procedure, probably not with complete success at the beginning, but showing a gradual improvement. It was this which the older subjects in particular could not do at the harder task; the shape of the error curve at Stage III (Figure 2) suggests that it was a difficulty which increased rapidly with age.

IV

Discussion

As this discussion does not aim to be exhaustive it will make no further reference to the behaviour of older subjects nor to the reasons why items which are spatially contiguous should be associated. But it will comment on two features, the Prediction of Mistakes and the Repetition of Errors, and precede this by a brief description of some pilot work to the experiment which has some bearing on these issues.

A Pilot Study.

In making preliminary observations the procedure was somewhat different. The subject began the task, as now, with the index card immediately beside the keys, but he was instructed that as he continued to do the task the card would gradually move away from the keys towards the lights. He was to ignore this change as far as possible and to carry on with his procedure no matter what the position of the card. This gave rise to some amusing and illuminating happenings. A subject would begin the task correctly, but as the card moved from the keys his responses would gradually take longer until, when the card was near the lights, he would perhaps stop altogether in some confusion. The experimenter would not say anything, but slowly bring the card back towards the keys, and as the card approached them, the subject would be able to resume successfully. It was a striking point that with some subjects this whole procedure could be repeated several times with just the same result and without the subject being able to analyse why he was able to do the task at some points and not at others.

It will be appreciated that this was a fascinating phenomenon to watch in so far as the difficulty of the operation could actually be measured on a linear scale. The "blocking" of the mental processes could be imposed or removed at will by sliding the card a matter of a few inches. It had been intended to present the experiment using this psycho-physical method, since it does establish this feature, but there was the objection that some subjects became so mystified by their failure that they suspected some form of "trickery" in the moving card.

Prediction of Mistakes.

The behaviour in the Pilot Study is, of course, an illuminating reiteration of that at Stage III, where the index card was immediately in front of the lights. At Stage III the difficulty was not so much one of an inability to apply an instruction because it was "forgotten," as an inability caused by one piece of information taking priority over another; that is in this case, the apparently contrary information from the spatial proximity of the lights and the index card overriding the original instructions. There are good grounds for saying that the failure was not one of "forgetting" the instructions, for the Type I errors, as already noted, are mainly caused by reversing the sequence of events—an indication that one stream of information was temporarily overriding or "blocking" another.

If subjects had been completely unable to do the task then it would have seemed, as it does in the problem-solving experiments of the "detour" type, that they were unable to think out the correct solution because they were too dominated by their interpretation of the perceptual situation confronting them. In the "detour" kind of experiments the assumption is generally made that the correct solution does not occur to a subject because there is one stream of evidence from the visual scene "suggesting" that the solution lies in one direction and apparently blinding a subject to an alternative. To solve the problem this misleading indication has to be overcome, and when this is achieved the experimenter has the difficulty of explaining what it is from a subject's previous experience, his mode of thinking, etc. which has aided him to achieve his solution.

In the present experiment whilst the evidence from the visual field—the external source—is comparable to that in the problem solving experiment, the information from the internal source has one important addition, namely, the immediately preceding experience of assimilating and executing the instructions for the task at the earlier Stage. Thus a subject is not searching, as he is in the Problem-Solving experiments, within his whole range of experience for a solution—he knows that this one lies within his immediate past. This is not attempting to deny the importance of a subject's general approach and mode of thinking, but emphasising the relevance for this task of one particular experience in it.

This kind of process whereby incoming data have to be reconciled with contrary instructions has often been described in terms of perceptual reorganisation. Here, on the basis of what we have discussed, it would seem that the procedure could be stated as follows: the organism was receiving data from the visual display. This input was subject to two particular kinds of modification and especially at Stage III the task was to reconcile these contrary indications.

(i) Items in a display which are spatially contiguous tend to be mentally associated. To do this would give rise to the two types of predicted errors and in the light of this occurence it can be claimed that there is phenomenal evidence how the input data were being treated.

(ii) Instructions had previously been assimilated that the perceptual data were not to be so treated, and a particular procedure was to be followed. The correct responses at Stage III indicate when this was achieved.

At the same time it should be noted that all subjects, without exception found, difficulty in carrying out Stage III; that is, they all took much longer than in the earlier stages and they nearly all made some errors. It would seem that the input had first come under the influence of the habitual mode of interpreting spatial contiguity (that is (i)); in other words the organism was first distorting its information before it tried to follow out the procedural instructions (ii). The experiment does not allow us

to be dogmatic on this point—it could be argued that the two procedures might be interacting simultaneously. But both the manner of subject's performances and the results would suggest that the temporal priority for these procedures was for data on their reception to be immediately interpreted by the organism according to its habitual modes of operation, and any analysis which was particular to conditions of the moment was subsequent to, and as here, often in opposition to the interpretation which had already been conducted. The whole problem may, in fact, be no more than this—a relatively simple analysis which has to break down the already established linkage of items. In this respect the experiment is a fair study of thinking, and it is of note that many subjects, after making some responses at Stage III correctly did, in fact, go on to make errors and become confused—it appeared that the established habitual procedure was able to assert itself when for any reason a hiatus occurred in a subject's thinking.

It is claimed then that the experiment in studying the adjustment between two contrary procedures and in being able to predict their respective results can be a little more precise about what are the exact difficulties which confront an organism in such a situation and how it has and has not adjusted to them.

Repetition of Mistakes.

The section in the results on repetition has brought out how frequently errors are repeated; at the risk of exhibiting the feature under discussion the following points seem worthy of emphasis.

We have argued that the mistakes are not a result of "forgetting" but of certain streams of information taking precedence over others—to a system of priorities. This is often a main difficulty of the problem-solving experiment, that the already held idea cannot be eradicated however often it has proved to be wrong and however much the subject himself is aware of its inadequacy. The effect then seems to be similar whether the mistaken idea arises from previous experience or from some feature in the immediate environment or some combination of both. The importance to accurate thinking of this ability to drop errors quickly has been generally recognised, in the field of mathematics notably by Hadamard (1945). From a slightly different standpoint Bartlett (1950) has considered how far subjects are prepared to continue in a course of actions against increasingly contrary indications. The present experiment partly fulfils some of his stipulations and throws some light on the question. The task has been so devised that the likely number of variants are few. If subjects continually do the task incorrectly they continue to repeat the same mistake though the evidence that this type of response is wrong is increasing; but equally they cannot move from one serial position to the next without following the right procedure and thereby receiving positive evidence on the point. The obvious reluctance, if not the disability of subjects to act upon this evidence, whether of the positive or negative variety is definite enough.

But the important issue for the writer is that this outstanding feature of problem-solving behaviour seems to be typical of what he observed in a rote learning situation with human subjects. In their rote learning there was the same reluctance by the poorer performers to accept information at a stage when it should have been patently obvious that the solution to which they were adhering—their already learned version—was wrong. Quick learners on the contrary had this ability to amend easily, to accept information at its face value and not to "argue" against it. It was this reluctance to discard their own ideas which characterised a poor performance both in the learning and in the present experiment. Not only were they puzzled by their mistakes, but n many instances there was a tendency, often overtly expressed, to argue that the

mistake was right, that the apparatus was faulty or something of that order. The frustrated reactions which were evoked were strikingly similar in both experiments. They were only to be solved by a definite act of "deliberation," in Johnson's use of the term (1944), a deliberation which contrary to his statement, was as much a characteristic of a rote-learning experiment as of the present problem. This seems reasonable enough—it is the same "biological" mechanism which is operating in either case and it is a feature of its functioning that new information has to be assimilated in relation to that which the system is already holding. Where instructions conform with habitual procedures they are quickly learned but, as here, the objectively same instructions may be only slowly acquired if they appear to contradict established habits.

Conclusion.

For the subjects who attempted this experiment it represented both a learning and a problem-solving task. They learned to carry out an operation, but they were also presented with a situation in which they had to think. As they had been given the solution they did not have to think out what that was, but rather had to consider how they were to apply it. Obviously this is not typical of all problem-solving experiments; here we have followed a restricted method by metaphorically turning the problem around and presenting a subject with the usual end product, the solution. This is very different from the more "intriguing" problems where the individual may have ranged over much of his imaginative and intellectual experience in searching for the solution. But it may well be that the more open type of problem is at present too ambitious for our experimental controls—we get some sort of result, but are too much in the dark how it has been obtained. The present technique does give us some precision which is open to further analysis. By predicting certain types of mistake it is possible to confirm what it is that is holding up a subject's progress. The error "solutions" which were put forward were those which had been considered and labelled. For a subject, progress was not barred by blankness but by wrong "solutions". For the experimenter, there was not a long and uninformative hiatus but ample opportunity to observe the interplay between the instructions and the changing spatial contiguities.

It is also felt that this use of the spatial dimension, more in the manner of the researcher on equipment design than in the traditional detour experiments is a useful precedent for problem studies. It need not be a barren model. The temporal dimension lends itself equally well to a similar experimental method, whilst other qualities such as size, shape or colour, could be treated on similar lines.

This study was carried out whilst the writer was a member of the Nuffield Research Unit into Problems of Ageing, Psychological Laboratory, Cambridge. It is desired to make acknowledgement to Messrs. Rowntree & Co. for their kindness in providing the facilities and the subjects for the experiment.

REFERENCES

BARTLETT, F. C. (1950). Programme for experiments on thinking. *Quart. J. exp. Psychol.*, **2**, 145–52.
HADAMARD, J. (1945). *The Psychology of Invention in the Mathematical Field*. Princeton.
JOHNSON, D. M. (1944). A modern account of problem-solving. *Psychol. Bull.*, **41**, 210–29.
KAY, H. (1951). Learning of a serial task by different age groups. *Quart. J. exp. Psychol.*, **3**, 166–83.
THORNDIKE, E. L. (*et al.*) (1928). *Adult Learning*. New York.
WELFORD, A. T. (1951). *Skill and Age*. London.
WELFORD, A. T. (1952). An apparatus for use in studying serial performance. *Amer. J. Psychol.*, **65**, 91–7.

SELECTIVE EFFECTS OF AN ANAESTHETIC DRUG ON COGNITIVE BEHAVIOUR

BY

HANNAH STEINBERG

From the Department of Pharmacology, University College London.

The prediction has been tested that under the influence of an anaesthetic drug, nitrous oxide, cognitive performances undergo differential impairment, the extent of which is positively correlated with the "complexity" of the task.

Ten kinds of performance were investigated, ranging from speed of finger tapping to reasoning by analogy. The relative complexity of each task was determined, in accordance with conventional criteria, from its respective qualitative category or "level"—relational, associative, and motor—and within each category from qualitative analysis of the component processes involved in its execution. A simple group difference design was used, involving two groups of 50 subjects each, matched for age and sex.

Significant deterioration as a consequence of drug administration occurred in the performance of all tasks. On the whole, the more complex a task the more did it tend to be impaired. Motor performances were, however, impaired to a greater extent than had been predicted. The possible significance of these findings is discussed.

INTRODUCTION

Behaviour can be modified by various drugs which act on the central nervous system, among them those which in higher doses produce anaesthesia. The precise mechanisms through which anaesthetic drugs act are imperfectly understood (Butler, 1950). Except for barbiturates, the effects of most of the well-known anaesthetics, including alcohol, ether, and nitrous oxide, are usually described following Hughlings Jackson's conception (1884) in terms of a law of "descending inhibition" (Harris, 1951; Himwich, 1951; Wilson & Schild, 1952). According to this the phylogenetically newer structures of the nervous system—the "higher centres"—mediating the most complex forms of behaviour, are depressed earlier and to a greater extent than the older "lower" structures.

Some psychologists have seen in this selective action corroboration for schemes of classifying "mental processes" according to their differential "complexity." Thus McDougall (1908, 1948) cited the allegedly progressive disorganisation of behaviour induced by alcohol as evidence for three mental "levels"—"intellectual," "sensory-motor" and "reflex." Recently Moursi (1952) has argued similarly in support of the "hierarchical" scheme developed from results of factor analyses by Burt (1949, 1949*a*), following which every cognitive process can "according to its relative complexity"—defined essentially in terms of correlation with intellectual ability—be assigned to one or other of four principal "levels" appropriate to "relational," "associational," "perceptuo-motor" and "sensory-motor" behaviour respectively. Diverse classifications of behaviour are, of course, permissible, and those like Burt's are useful mainly because they correspond to familiar qualitative categories. Nevertheless, agreement between classifications developed from two such different criteria as a form of psychological analysis on the one hand and the selective action of drugs on the other, would contribute to the understanding of both.

The assumptions behind the law of descending inhibition require modification in the light of contemporary views on relations between the nervous system and behaviour. The evidence is in most respects still inconclusive and, as is well known, there can be marked differences between effects of different anaesthetics. However,

it is at any rate broadly probable that there is a positive correlation between complexity of interneuronal organisation in the nervous system and complexity of overt behaviour, and also between complexity of interneuronal organisation and susceptibility to the action of anaesthetics (Bárány, 1947; French *et al.*, 1953). From this it should still follow that behaviour will be impaired selectively and according to its relative complexity by a given amount of anaesthetic. Relevant findings by previous investigators, though numerous especially on alcohol, are difficult to interpret because of divergencies in design and techniques and because most of them demonstrate the presence or absence of an effect on individual kinds of performance without comparable measurement of the magnitude of disturbance. Most findings seem not incompatible with expectation (e.g. Varner, 1933; Gantt, 1935; Newman *et al.*, 1942; Lauer, 1939; Mead, 1939; Jellinek & McFarland, 1940; Spragg, 1941; Wechsler, 1941; Masserman, 1945, all on alcohol; McKinney, 1932; Marshall, 1937, on nitrous oxide; Case & Haldane, 1941, on "nitrogen narcosis" attendant upon increased air pressure). Some conflict (e.g. Seward & Seward, 1936; Davis *et al.*, 1941, both on alcohol). In particular, the relative susceptibility to anaesthetics of relational and associative behaviour on the one hand and of motor behaviour on the other has not been adequately explored.

The main purpose of the present investigation was to examine the effects, which have been little studied hitherto, of the anaesthetic gas nitrous oxide, in a concentration not involving loss of consciousness, on several aspects of cognitive behaviour. It was predicted that (i) under the influence of nitrous oxide cognitive behaviour tends to deteriorate, and (ii) this deterioration occurs selectively; the more "complex" a performance the more it is impaired. A brief preliminary account of some of this work has already been published (Steinberg, 1951); fuller details are contained in a thesis by the author (Steinberg, 1953).

Method and Procedure

I. Experimental Design.

The method of difference design used is illustrated in Table I. One hundred subjects, all university students, were divided into two equal groups matched for age and sex. Each subject attended on two occasions, separated by one week. On the first occasion treatment was identical for members of both groups, measures of performance in 10 different cognitive tasks or "tests" being obtained from each subject. On the second occasion measures

TABLE I

Experimental Design

Group	*Treatment**	
	Trial I	*Trial II*
Experimental 50 subjects	10 cognitive tasks No Drug	10 cognitive tasks Drug
Control 50 subjects	10 cognitive tasks No Drug	10 cognitive tasks No Drug

* Trials I and II were separated by one week.

of performance in 10 tasks similar to those given earlier were again obtained from all subjects; but experimental subjects inhaled the drug throughout this phase, while control subjects inhaled air. Thus the presence or absence of the drug is the independent variable, and behaviour in the test series constitutes the dependent variable. If it can be shown that the two groups are equated on initial performance, comparison of their

subsequent performance enables the drug effect to be measured. Finally, the experimental design makes possible comparisons between the different kinds of cognitive behaviour sampled in respect of any differential changes undergone by them under the influence of the drug.

II. Administration of the gases.

Nitrous oxide (N_2O) is an anaesthetic gas which, administered by inhalation, is rapidly absorbed and excreted through the lungs. The apparatus used enabled administration to be maintained at a relatively constant rate for any desired period of time. The flow of gases was measured by rotameter-type flow-meters and delivered by an open circuit to a reservoir bag of 3½ gallons capacity connected by corrugated rubber tubing to a modified R.A.F. oxygen mask. A microphone in the mask enabled the subject to communicate with the experimenter. The apparatus delivered 30 per cent. nitrous oxide in oxygen at a rate of 15 litres per minute to all experimental subjects, this concentration being chosen as preliminary experiments had shown it to induce marked changes in most subjects' cognitive behaviour without seriously impairing their co-operativeness. The air inhaled by the control subjects was administered from the same apparatus and at the same rate. Any unused surplus escaped through expiratory valves in the mask. Since nitrous oxide normally has a distinctive smell, all gases were perfumed by inserting into the circuit a filter-paper impregnated with lavender oil. Five minutes' preliminary inhalation was always given prior to the testing session proper, since this period has been shown to be adequate for establishing a relatively stable equilibrium between the concentrations of nitrous oxide in the arterial and venous blood supply of the brain (Kety & Schmidt, 1948).

III. Tests of behaviour.

The tests chosen cover a wide range of cognitive behaviour, taking into consideration the practical requirements of brevity, standardisation, repeatability, and convenient administration to subjects hampered by wearing somewhat cumbersome apparatus. The following 10 tests were used:—

A. Relational.

1. "Non-Verbal Analogies": a multiple choice test containing 30 diagrammatic items arranged in order of increasing difficulty. Score: number of items correctly answered in 2½ minutes. This test and test 2 were adapted from standardised material kindly supplied by Dr. Edgar Anstey.

2. "Verbal Analogies": a similar test, but consisting entirely of verbal material.

B. Associative.

3. "Arithmetic": the subject added as quickly as possible sums of four 2-digit numbers. Score: number of individual columns correctly added in 2 minutes.

4. "Digit-span Backwards": the subject repeated in reverse order sequences of digits of increasing length presented to him orally at the rate of one digit per second. Score: number of digits in longest sequence repeated in correct order, two trials being allowed for each length of sequence.

5. "Fluency 'Flowers' ": the subject said as many names of flowers as possible in the time allowed. Score: number of names of flowers given in 60 seconds. (Eysenck, 1947.)

6. "Fluency 'Things to Eat' ": as 5 above, but calling for names of things to eat. (Eysenck, 1947.)

7. "Digit-span Forwards": as 4 above, except that the digits were to be repeated in the same order as presented.

C. Motor.

8. "Dotting": McDougall-Schuster Disc Dotting Machine embodying a revolving spiral of irregularly placed dots at which the subject aimed with a pencil. Score: number of dots "hit" before two consecutive misses.

9. "Ball-bearing": the subject inserted with a pair of forceps steel balls as quickly as possible into a vertical tube. Score: number of balls inserted in 40 seconds. (The apparatus resembled the R-V manual dexterity test developed by the National Institute of Industrial Psychology.)

10. "Tapping": the subject tapped with his hand as quickly as possible a Morse key connected to an electric counter. Score: number of taps in 30 seconds.

Parallel versions were used for tests 1, 2, 3, 4 and 7 in the two presentations.

The tests have been listed above in the order of their relative "complexity". For the purposes of the present experiment this was determined for each test first from its respective qualitative category or "level," i.e. relational, associative, and motor. This classification corresponds to Burt's, except that his two "lowest" categories have been combined into one. Individual tests in each category were then ranked by qualitative analysis of the component processes involved in their execution. Thus, for example, "Non-Verbal Analogies" ranked higher than "Verbal Analogies" on account of its less familiar content—diagrams as against words—and both these tests, being "relational," ranked higher than all the remaining tests. The complete rank order was determined unanimously by three judges working independently. Some further suggestive corroboration was obtainable for the order assigned. First, Moursi (1952) ranked four of his tests "Non-Verbal Analogies," "Verbal Analogies," "Verbal Fluency," and "Digit-Span Forwards"—which closely resemble those used in the present investigation, in the same order. His ranking was based on the relative saturation in the general factor. Secondly, where tests in the present experiment were scored along comparable scales, the relative size of the mean scores in the first trial is a possible index of differential complexity, and, as can be seen later from Table II, is compatible with the ranks already assigned. For example, "Digit-Span Backwards" gave rise to a lower mean score than "Digit-Span Forwards," and "Fluency 'Flowers'" to a lower score than "Fluency 'Things to Eat'".

Results

Table II summarizes the results, in raw scores, obtained by the two groups of subjects in the ten performances tested during both phases of the experiment.

In order to test the two predictions under investigation, these results were analysed as follows:—

(i) Matching of experimental and control groups in respect of initial test performance.
(ii) Overall effect of the drug on test performance.
(iii) Differential effects of the drug on test performances.
(iv) Comparison of observed and expected differential drug effects.

In addition, so as to obtain measures of reliability of the tests used, correlations were calculated between the test scores obtained in the first and second trials by the control group. The resulting coefficients ranged from $+ 0{\cdot}573$ to $+ 0{\cdot}888$ and were all statistically significant at the 0·001 level.

(i) *Matching of experimental and control groups in respect of initial test performance:*

To test the comparability of the two groups in respect of initial performance, Hotelling's T test using "paired scores" was used (Hotelling, 1931), the null hypothesis being that, assuming equal variances and co-variances, there was no difference between the two groups in respect of mean scores on all ten tests. T^2 was found to be 13·11, and the corresponding F ratio 0·95, which is not statistically significant ($P > 0{\cdot}2$). Hence the null hypothesis was accepted and the two groups were regarded as satisfactorily equated on initial performance.

(ii) *Overall effect of the drug on test performance:*

To take into account all the information available, both in respect of the matching of individuals between the two groups and of the testing and retesting of each individual in each group, the calculation of the drug effect was based on "second-order differences." For each subject ten difference scores were calculated by subtracting his first score from his second score on each test. Then for each matched pair of subjects ten second-order difference scores were calculated by subtracting the difference scores of the control subject from those of the experimental subject. Hotelling's T test was used to test the overall statistical significance of these second-order differences, the null hypothesis tested being that the drug had introduced no

TABLE II

SUMMARY OF TEST RESULTS IN RAW SCORES

Test	Parameter	Experimental N = 50			Control N = 50			2nd order diff.
		Trial I	Trial II	Diff.	Trial I	Trial II	Diff.	
1. Non-Verbal Analogies	Mean s.d.	10·50 2·40	8·22 2·77	− 2·28	9·64 2·55	11·06 2·21	+ 1·42	− 3·70
2. Verbal Analogies	Mean s.d.	15·10 4·51	13·36 4·13	− 1·74	14·24 3·77	18·80 3·78	+ 4·56	− 6·30
3. Arithmetic	Mean s.d.	23·46 7·40	14·94 7·23	− 8·52	24·84 7·00	24·42 7·71	− 0·42	− 8·10
4. Digit-Span Backwards	Mean s.d.	5·66 1·53	4·62 1·35	− 1·04	6·06 1·39	6·52 1·43	+ 0·46	− 1·50
5. Fluency "Flowers"	Mean s.d.	18·04 5·52	15·42 5·72	− 2·62	17·20 4·98	19·72 5·50	+ 2·52	− 5·14
6. Fluency "Things to Eat"	Mean s.d.	28·50 7·11	23·30 7·48	− 5·20	28·18 7·68	28·44 8·28	+ 0·26	− 5·46
7. Digit-Span Forwards	Mean s.d.	7·58 1·56	7·02 1·45	− 0·56	7·78 1·24	8·00 1·13	+ 0·22	− 0·78
8. Dotting	Mean s.d.	151·24 44·82	109·10 36·89	−42·14	168·06 46·83	178·18 52·06	+10·12	−52·26
9. Ball-Bearing	Mean s.d.	16·00 1·91	13·12 2·04	− 2·88	16·52 2·29	17·02 2·51	+ 0·50	− 3·38
10. Tapping	Mean s.d.	198·48 25·28	172·32 20·53	− 26·16	201·34 22·84	199·36 22·06	− 1·98	− 24·18

difference in the performance of the experimental group, as compared with the control group. The T^2 ratio was 35·94, and the corresponding F ratio 25·47 which, with d.f. 10 and 40 is highly significant statistically ($P < 0·001$). Hence the drug had been shown to have had a significant adverse effect on test performance.

In addition, individual t ratios, using the formula for "paired scores," were calculated and showed that the mean second-order differences in respect of each of the ten tests were statistically significant at the 0·005 one-tail level, or higher. It is of interest also to note from inspection of the standard deviations given in Table II that variability of performance appears to have been fairly uniform throughout, including that of the experimental group when the drug was being given.

TABLE III

DIFFERENTIAL IMPAIRMENT OF PERFORMANCES ATTRIBUTABLE TO DRUG

Tests (in rank order of expected impairment)	*Observed impairment*	
	s.d. difference score	*Rank order*
1. Non-Verbal Analogies ..	1·471	3
2. Verbal Analogies	1·508	2
3. Arithmetic	1·120	5
4. Digit-Span Backwards ..	1·016	6
5. Fluency "Flowers"	0·975	8
6. Fluency "Things to Eat" ..	0·738	9
7. Digit-Span Forwards ..	0·552	10
8. Dotting	1·121	4
9. Ball-Bearing	1·591	1
10. Tapping	1·002	7

It will have been realized that for the purpose of these experiments "deterioration of performance" is being defined "operationally" simply as a reduction in test scores, and that the test scores are in all cases based on quantity of correct output in a given time and/or a given number of attempts. Hence a drop in score is equivalent to a decrease in efficiency. For the three written tests—"Non-Verbal Analogies," "Verbal Analogies," "Arithmetic"—it was, however, also possible to examine whether the drop in score was attributable primarily to a decrease in the total number of test items attempted, to an increase in the number of errors, or to both. The statistical significance of the differences between the experimental and control groups in respect of all relevant pairs of distributions of these two variables was calculated by means of χ^2 tests using medians (Mood, 1950). This method of calculation was adopted because of the pronounced skewness of some of the distributions. The results showed that neither for the total number of items attempted nor for the number of errors were the differences between the experimental and control groups statistically significant in the

first trial, but they were significant, at the 0·01 level or higher, in the second trial in respect of all three cognitive performances considered. This suggests that both factors contributed to the reduction in test scores during inhalation of the drug.

(iii) *Differential effects of the drug on test performances:*

To test the second prediction, that deterioration of performance occurs selectively and in accordance with the relative complexity of the performances, the test scores from different tests had first to be made comparable. The second-order difference scores were accordingly converted into standard deviation difference scores, i.e. the ratio was computed of the mean second-order difference to the standard deviation of test scores obtained by both groups in the first trial. These standard deviation scores could be used as measures of the effects of the drug, and the different tests arranged in rank order accordingly, from most to least impaired. The results are shown in Table III and illustrated in Figure 1. It will be seen that the relational tests were impaired to a greater extent than associative tests, but that performance in motor tests was also much affected, the Ball-Bearing test appearing the most sensitive of all.

FIGURE 1

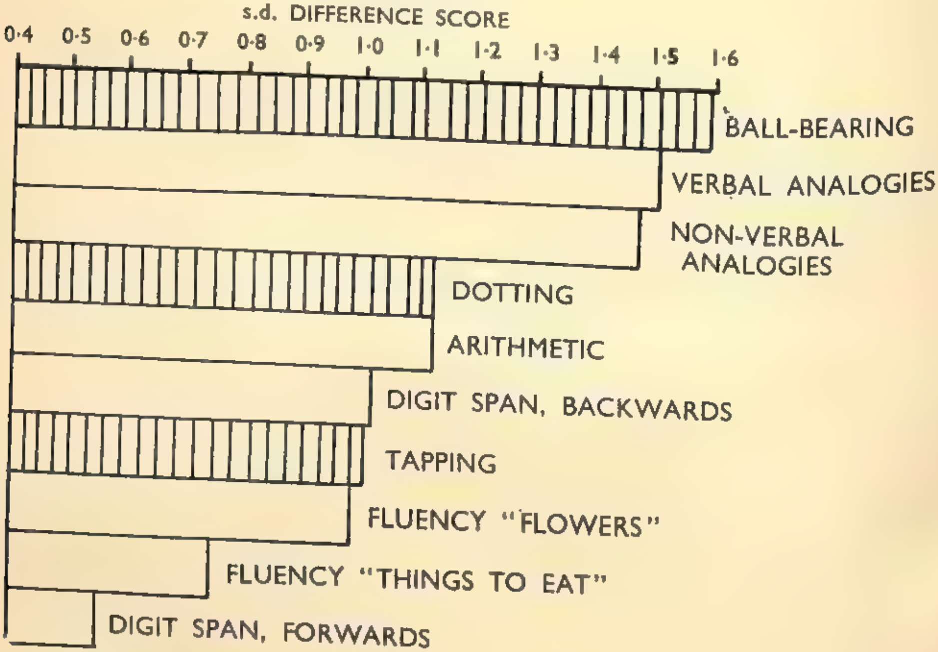


Diagram of differential impairment of performances attributable to drug.

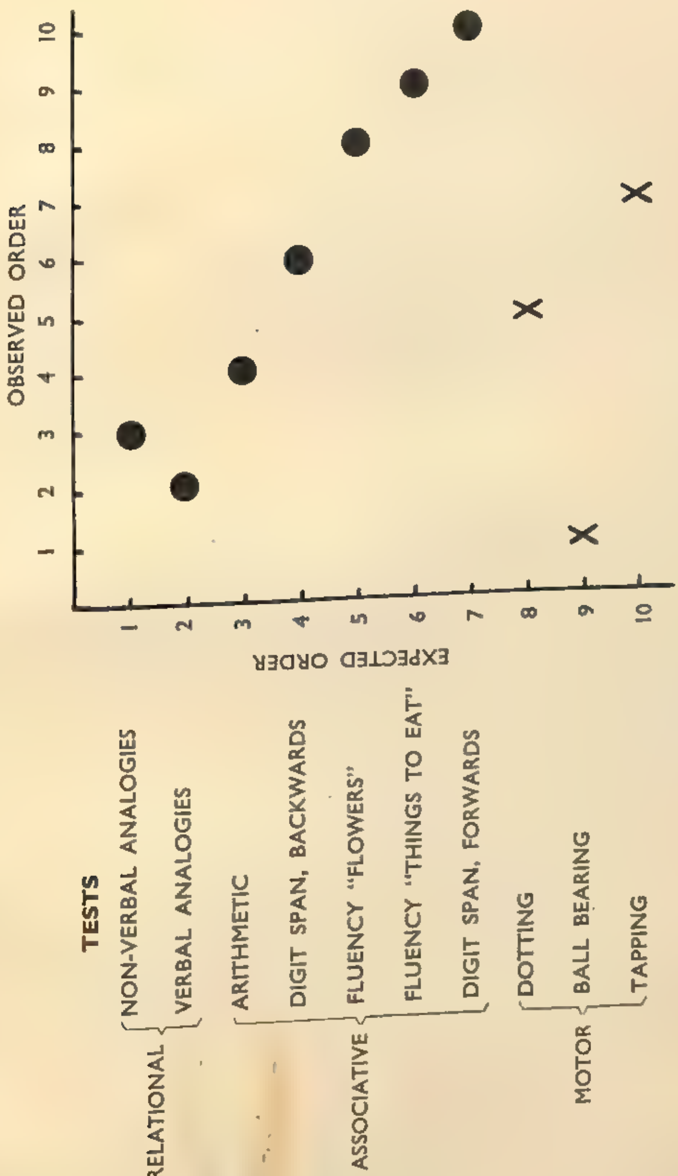
It may be noted that the method of conversion into standard scores described does not enable the statistical significance of the differential impairment between the different tests to be estimated, nor does it take account of correlations between the tests in respect of second-order differences. As the calculated values of these correlations were found to be almost uniformly small it seemed reasonable to omit such a refinement.

(iv) *Comparison of observed and expected differential drug effects:*

The final step in testing the second prediction was to compare the differential drug effects actually found with those expected on the basis of the differential

"complexity" of the performances tested. As has been explained, the tests used had been ranked from test 1, the most complex, to test 10, the least complex. Hence differential impairment of performances under the influence of the drug was expected to follow the same order, test 1 being expected to be most impaired and test 10 least.

FIGURE 2



Relation between expected and observed orders of impairment

The observed rank order of impairment compared with the expected order gives a Spearman rank correlation of + 0·285. This fails to reach statistical significance ($P > 0{\cdot}05$). It will be noted from the dot diagram in Figure 2, however, that the discrepancies between the expected and observed rank order are due almost entirely

to the motor tests. If these are eliminated from the rank orders, the correlation rises considerably, to + 0·964 (P = 0·0014). These results suggest that the prediction has been substantially verified as far as the relational and associative tests are concerned; the motor tests, however, appear to have been impaired to a different and greater extent than was predicted in terms of the criteria of relative complexity here adopted.

Discussion and Conclusions

The results described have the following implications:

(*a*) The overall efficiency of cognitive performance deteriorates significantly under the influence of a sub-anaesthetic concentration of nitrous oxide. It has, in addition, been shown that there was significant impairment of performance on each of the individual cognitive tasks investigated, and that in the three written tasks this impairment took the form both of decreased total number of items attempted and of increases in the number of errors. Of particular incidental interest is the definite decline which occurred in verbal fluency, in view of the fact that nitrous oxide and similar drugs are often used in psychotherapy with the object of making the patient more loquacious (Horsley, 1943). It should be pointed out that it is not possible on the basis of the present investigation to decide how far deterioration of performances was due to a lowering of cognitive ability as such, and how far to such factors as reduced motivation, impaired concentration, or interference from various physical and psychological symptoms experienced during drug administration. However, there is no correlation between individual differences in the frequency with which these latter were reported and individual differences in measures of impairment of cognitive performance. This suggests that the symptoms did not have any systematic effect on cognitive performance. In three of the tests, as has been described, significant increases occurred in the number of items attempted but answered wrongly. Thus there is a possibility that although subjects under the influence of the drug were less able to make correct solutions, they may have become more willing to guess—an interpretation supported by comments made by several subjects.

(*b*) In the case of relational and associative performances, the relative amount of deterioration is closely related to the "complexity" of the task, as determined from an analysis of qualitative characteristics. The more complex a performance, the more it tends to deteriorate. The correlation between the expected and obtained rank orders of impairment, excluding the motor tests, was very high.

(*c*) Motor performances are, however, capable of being impaired to a greater extent than was predicted in terms of the definition of complexity adopted. Performance in all three motor tests used was more impaired than performance in some of the associative tests; and the Ball-Bearing test was affected to a greater extent than all other tasks. Detailed generalizations cannot justifiably be based on results from only three tests. Nevertheless, inspection of the data does perhaps suggest that relatively complex motor tasks, such as performance in the Ball-Bearing or Dotting tests, may tend to be more susceptible to the effects of an anaesthetic drug than very simple associative tasks, such as repeating digits forwards. Difficult associative tasks, on the other hand, may tend to be more sensitive than very simple motor ones.

The present investigation therefore suggests the tentative conclusion—which is not incompatible with some of the findings by other workers—that where the effects of anaesthetic drugs are concerned, relational and associative tasks on the one hand, and motor tasks on the other, should be considered separately. Within either category, the more complex the task, the more it is likely to be impaired. The evidence from the present experiments is not, however, consistent with the assumption that all relational and associative tasks are more complex than motor tasks, the view implicit

in the classificatory schemes of McDougall, Burt, and Moursi. From a neurophysiological point of view, a possible interpretation is that the assumption that anaesthetic drugs affect behaviour strictly in accordance with the relative complexity of the corresponding interneuronal pathways may not be true in the present case, but that those pathways primarily concerned with motor behaviour are especially sensitive to the drug used. An alternative possibility would appear to be that the complexity of neuronal organization involved in symbolic behaviour is not invariably greater than in all forms of motor activity.

Acknowledgements

I am most grateful to Professor F. R. Winton for constant encouragement and help with all aspects of this work; to Professor R. W. Russell for much ready help and interest; to Professor Sir Cyril Burt for valuable advice during the earlier stages; to Mr. A. Summerfield for advice on statistical and many other problems; to Mr. J. W. Whitfield, Dr. J. A. B. Gray, and Mr. J. Diamond for helpful criticism of the manuscript; to members of University College London, who acted as subjects; to Group-Captain L. K. Stewart, Commanding Officer of the R.A.F. Institute of Aviation Medicine, for the loan of oxygen masks; and to the Medical Research Council for financial assistance.

References

Bárány, E. H. (1947). A theoretical note concerning the action of drugs on the central nervous system. *Arch. int. Pharmacodyn.*, **75**, 222–226.

Burt, C. (1949). The structure of the mind. *Brit. J. educ. Psychol.*, **19**, 100–111, 176–199.

Burt, C. (1949*a*). Subdivided factors. *Brit. J. Psychol.* (Stats. Section.), **2**, 41–63.

Butler, T. C. (1950). Theories of general anaesthesia. *Pharmacol. Rev.*, **2**, 121–160.

Case, E. M. and Haldane, J. B. S. (1941). Human physiology under high pressure. I. Effects of nitrogen, carbon dioxide and cold. *J. Hyg., Camb.*, **41**, 225–249.

Davis, P. A., Gibbs, F. A., Davis, H., Jetter, W. W. and Trowbridge, L. S. (1941). The effect of alcohol upon electro-encephalogram. *Quart. J. Stud. Alc.*, **1**, 626–637.

Eysenck, H. J. (1947). *Dimensions of Personality.* London.

French, J. D., Verzeano, M. and Magoun, H. W. (1953). A neural basis of the anaesthetic state. *Arch. Neurol. Psychiat., Chicago.*, **69**, 519–529.

Gantt, W. H. (1935). Effect of alcohol on cortical and subcortical activity measured by the conditioned reflex method. *Bull. Johns Hopkins Hosp.*, **56**, 61–83.

Harris, T. A. B. (1951). *The Mode of Action of Anaesthetics.* Edinburgh.

Himwich, H. E. (1951). *Brain Metabolism and Cerebral Disorders.* Baltimore.

Horsley, J. S. (1943). *Narco-analysis.* London.

Hotelling, H. (1931). The generalisation of student's ratio. *Ann. math. Stat.*, **2**, 360–378.

Jackson, Hughlings (1884). The Croonian Lectures on Evolution and Dissolution of the Nervous System. *Brit. med. J.*, **1**, 591–599; 660–663; 703–707.

Jellinek, E. M. and McFarland, Ross, A. (1940). Analysis of psychological experiments on the effects of alcohol. *Quart. J. Stud. Alc.*, **1**, 272–371.

Kety, S. S. and Schmidt, C. F. (1948). The nitrous oxide method for the quantitative determination of cerebral blood flow in man. *J. clin. Invest.*, **27**, 476–483.

Lauer, A. R. (1939). The effects of alcohol on driving. *J. Iowa State med. Soc.* **29**, 282–285.

McDougall, W. (1908). *Physiological Psychology.* London.

McDougall, W. (1948). *An Outline of Abnormal Psychology.* London.

McKinney, F. (1932). Nitrous oxide anaesthesia as an experimental technique in psychology. *J. gen. Psychol.*, **6**, 195–199.

Marshall, C. R. (1937). The influences of moderate and severe intoxication on remembering. *Brit. J. Psychol.*, **28**, 18–27.

Masserman, J. H. (1945). Neuroses and alcohol. *Amer. J. Psychiat.*, **101**, 389–397.

Mead, L. C. (1939). The effect of alcohol on two performances of different intellectual complexity. *J. gen. Psychol.*, **21**, 3–23.

Mood, A. M. (1950). *Introduction to the Theory of Statistics.* New York.

MOURSI, E. M. K. M. (1952). A factorial investigation of the hierarchical organisation of cognitive levels. Ph.D. thesis. University of London.

NEWMAN, H. W., FLETCHER, E. and ABRAMSON, M. (1942). Alcohol and driving. *Quart. J. Stud. Alc.*, **3**, 15–30.

SEWARD, G. H. and SEWARD, JOHN P. (1936). Alcohol and task complexity. *Arch. Psychol.*, **29**, 5–26, *N.Y.*

SPRAGG, S. D. S. (1941). The effects of certain drugs on mental and motor efficiency. *Psychol. Bull.*, **38**, 354–363.

STEINBERG, H. (1951). Some effects of depressant drugs on behaviour. *Proc. XIII International Congress of Psychology, Stockholm*, 221–222.

STEINBERG, H. (1953). Some effects of depressant drugs on behaviour. Ph.D. thesis. University of London.

VARNER, W. B. (1933). The effects of alcohol on two maze habits of albino rats. *Psychol. Bull.*, **30**, 616.

WECHSLER, DAVID (1941). The effect of alcohol on mental activity. *Quart. J. Stud. Alc.*, **2**, 479–485.

WILSON, A. and SCHILD, H. O. (1952). Clark's *Applied Pharmacology*. London.

DRIVE-LEVEL AND FLEXIBILITY IN PAIRED-ASSOCIATE NONSENSE-SYLLABLE LEARNING

BY

WALTER M. GORDON and D. E. BERLYNE

*From the University of Aberdeen**

"High-anxiety" and "low-anxiety" subjects, selected for extreme scores on the Taylor Anxiety Scale, learned a list of paired-associate nonsense syllables in the belief that they were undergoing an intelligence test. Both groups were then given a second list of paired associates to learn, the stimulus-items being the same as those of the first list but the responses being changed. Before the presentation of the second list, half the subjects in each group were given anxiety-increasing instructions and the remaining half were given reassuring instructions.

The results verified two predictions made from Hull's behaviour theory, using the concept of fear or anxiety as a secondary drive:—"high anxiety" subjects took more trials to master the second learning task than "low-anxiety" subjects; and there was a significant interaction between initial anxiety-level and type of instructions, such that "high-anxiety" subjects who received drive-increasing instructions had a worse performance in the second part than all other sub-groups. There was no indication that "low-anxiety" subjects were significantly affected by the type of instructions received. The "high-anxiety" group had greater difficulty than the "low-anxiety" group in learning the first list, but the difference was non-significant.

I

Introduction

An increasingly popular field for experiment in recent years has been response fixation. Fixation may be defined as difficulty in abandoning a response when it is no longer adaptive, and the questions it raises are not only of intrinsic theoretical interest but also highly relevant to the problem of human neurosis, which usually involves a strange inability to give up forms of behaviour that are clearly punishing.

Some writers (e.g. Maier, 1949) have asserted that fixation resulting from frustration is a unique phenomenon which requires a special principle to account for it. But others (e.g. Farber, 1948; Mowrer, 1950; Wilcoxon, 1952) have offered arguments and experimental data pointing in a different direction. According to this other view, fixation can be explained by familiar principles of S–R behaviour theory, including the assumption that fixated responses are motivated by the secondary drive of fear and reinforced by fear-reduction.

Another relevant series of studies has been concerned with the effects of anxiety-level on performance as distinct from the effects of anxiety-reduction on learning. Spence and his collaborators have selected subjects according to their scores on the Taylor Anxiety Scale (Taylor, 1951, 1953). What exactly the Taylor Scale measures has been a matter for some controversy, but Spence's team have succeeded in verifying predictions from Hull's behaviour theory (Hull, 1943, 1952) with the assumption that a high score on the scale reflects a high drive-level (D). The predictions and experiments (Spence & Taylor, 1951; Taylor, 1951; Spence & Farber, 1953) indicate that "high-anxiety" subjects acquire a simple classical conditioned response more rapidly than others. But in complex learning (Taylor & Spence, 1952; Montague, 1953; Farber & Spence, 1953), i.e. when there are response-tendencies in competition

* Based on a thesis submitted by W. M. Gordon in accordance with the requirements for the degree of M.A. with Honours in Psychology of the University of Aberdeen.

with the response that is to be mastered, "high-anxiety" subjects are at a disadvantage. Since competition between responses is clearly present when one form of behaviour has to be replaced by another, the relation between anxiety-level and performance may provide a further clue to cases of fixation or lack of flexibility.

The present paper reports an experiment intended to investigate this problem with paired-associate nonsense-syllable learning in human beings. Our subjects had first to learn a response syllable to each of a set of stimulus-syllables ($S_1 \rightarrow R_1$). They had then to learn different response-syllables ($S_1 \rightarrow R_2$) to the same stimuli. The degree of difficulty in the second task is not, strictly speaking, identifiable with fixation, since it does not imply persistence of the overt responses appropriate to the first task. But insofar as the learning of the second series of responses is impeded by the effects of learning the first, we have an essentially similar situation. Our subjects were divided into four sub-groups. First of all, "high-anxiety" and "low-anxiety" individuals were selected by the Taylor Anxiety Scale as modified by Cutts (1953). Then, both "high-anxiety" and "low-anxiety" subjects were further subdivided according to the treatment received between the first and second phases of the experiment. Before attacking the second task, "experimental" sub-groups were given instructions designed to raise their anxiety-levels, and "control" sub-groups were given instructions designed to minimize their anxiety.

Our experiment tests predictions from Hull's behaviour theory as follows. (It will be noted that our reasoning parallels that used by Spence and his collaborators with reference to complex learning).

1. It is assumed that subjects with high scores on the anxiety scale have a higher drive-level (D) than those not scoring highly.

2. An increase in D will result in an increase in reaction-potential [sEr = f(sHr) x f(D)]. (This is Postulate 7 in Hull's earlier work (1943). In his last book (1952), it is further elaborated and becomes Postulate VIII.)

3. It is further assumed that probability of response evocation increases with reaction-potential. (This is Postulate 12 (Hull, 1943), and it appears to be assumed without being explicitly stated in the posthumous book (1952).)

4. An increase in D will increase the strength of all response-tendencies in the organism. If initially one such tendency is stronger than another, an increase in D will increase the advantage of that particular tendency (da -db = d (a–b) where "d" is a function of D and "a" and "b" are functions of the two response-tendencies respectively). Thus, if a subject has to substitute a new (initially weak) response for an old and stronger response, the former, before it becomes dominant, will have more "leeway" to make up.

5. (*a*) It follows that subjects with higher scores on the anxiety questionnaire will find greater difficulty in abandoning one response to a stimulus in favour of another, i.e. such subjects will require a greater number of trials to learn a new list of responses.

(*b*) Moreover, the stress-reducing instructions will decrease the disadvantage of the "high-anxiety" subjects and the stress-increasing instructions will exacerbate it. We thus predict an interaction between the variables, such that the greatest difficulty will occur with the experimental "high-anxiety" subjects.

II

The Experiment

Subjects:

The subjects were selected from the introductory class in psychology on the basis of their scores on the Taylor Anxiety Scale (Taylor 1951, 1953) as modified by Cutts (1953).

The "high-anxiety" and "low-anxiety" groups each consisted of 20 students whose scores fell respectively within the highest 32 per cent. and lowest 37 per cent. of scores for a population of 83 students. All subjects were unaware of the reason for their selection. Their mean age was 19·4 years, the mean ages for "high-anxiety" and "low-anxiety" groups being 19·5 and 19·3 respectively. The "high-anxiety" group contained 14 female subjects as compared with 10 in the "low-anxiety" group.

Apparatus:

The apparatus used for presenting the nonsense syllables was the Ranschburg memory drum, an illustration of which is shown by Katz (1948, p. 63). The time-pulse required to operate the apparatus was obtained from 3 two-volt accumulators, connected in series via a Palmer timing-unit.

Materials:

Fifteen three-letter nonsense syllables were chosen from the 0 per cent. and 7 per cent. association-value lists calibrated by Glaze and reproduced by Hilgard (1951). These were arranged into two lists of paired associates each consisting of 5 pairs—the 5 stimulus items being the same for each list. The 5 responses of the first list included 3 items of 0 per cent. association value and there were 4 items of this value in the second response list. In compiling the two lists an attempt was made, as far as possible, to conform to the rules set out by Hilgard (1951, p. 540). The stimulus items (the first members of the pairs) were marked with a red spot to help the subject to distinguish them from the responses.

Each syllable was exposed for 2 seconds, and there was a 2-second interval between each stimulus item and its associated response and between that response and the next stimulus item. Three orders were used for each of the two lists to reduce serial learning as far as possible. This follows a common practice since McGeoch and McKinney (1937) found that it is easier to learn paired associates if the order remains constant than if it is varied. The lists were each employed in 3 permutations using tables of random numbers, and these 3 permutations were used in arranging the syllables on the discs, so that for each list the different permutations always followed each other in the same sequence. A practice list containing 6 paired items of higher association value than those used in the actual experiment was presented in the apparatus once to all subjects.

Procedure:

Before the experiment proper began the subject was asked for his co-operation and requested to refrain from asking questions until the experiment was over, apart from making certain that the instructions were clearly understood. The subject was informed that he was to perform a new kind of intelligence test of greater validity than most, and he was then given information about the method of paired-associate learning. In particular, the subject had to spell out each syllable and try to anticipate the response syllables by spelling them out before they appeared in the aperture. All subjects were instructed to guess if they thought they knew a particular response item but were not sure of it. They were also told that learning would continue until they had reached the criterion of two successive errorless trials. In an attempt to ensure that the instructions had been clearly understood, subjects were given one presentation of the practice list and they could ask questions at this point if they were in doubt.

The subject then performed the first part of the experiment, having rest pauses of 30 seconds after every 15 trials (i.e. every 10 minutes). The experimenter noted both correct and incorrect responses and the number of trials required to reach the criterion. The subject then rested for 2 minutes, during which time the experimenter made a show of swiftly calculating the score obtained. Then the disc containing the new pairs was inserted in the apparatus behind the screen. Allocation to the "experimental" or the "control" treatment was by alternation, i.e. a subject was included in the experimental group if the previous one had been in the control group and *vice versa*. Subjects in the experimental sub-groups were told that their scores were below the average for university students and that they would now have a second run so that they might redeem themselves. Control-group subjects heard that they were above average for students and that they would now have a second run designed to give information on exactly how much above average they were. All sub-groups were told that the stimulus-syllables were the same as before but that the responses were new; of the two lists used, each was presented first and the other second for half the subjects in each sub-group.

III

Results

Part I. An analysis of variance was carried out on the numbers of trials to reach the criterion for the first learning task. The means were 31·25 for the "high-anxiety" subjects and 26·25 for the "low-anxiety" subjects, but no significant F emerged either for this comparison or for any other. Both Bartlett's test for homogeneity of variance and the F-test failed to show any significant differences in variance between "high-anxiety" and "low-anxiety" groups.

Part II. The means for trials to reach the criterion in the second phase are displayed in Table I. It will be seen that the experimental "high-anxiety" sub-group took longer to learn than the other three sub-groups but that the means of the other three sub-groups are closely comparable. All sub-groups, however, showed an improvement on the means for Part I.

An analysis of variance produced significant Fs for two comparisons, viz. "high-anxiety" versus "low-anxiety" subjects ($F=6{\cdot}75$, $df=1,32$, $p<{\cdot}05$) and interaction between initial anxiety-level ("high-anxiety" versus "low-anxiety") and instructions (experimental versus control) ($F=5{\cdot}31$ $df=1,32$, $p<{\cdot}05$). The difference between experimental and control groups was not shown to be significant.

Since there was a difference, though not a significant one, between the "high-anxiety" and "low-anxiety" groups in part I, it was judged worthwhile to perform an analysis of covariance on the data from both phases of the experiment. This analysis revealed once again an anxiety-level x instructions interaction that is significant at the ·05-level ($F=4{\cdot}55$) But the F for anxiety-level this time fails to reach significance. Thus, we are forced to consider the possibility that the results for the "high-anxiety" and "low-anxiety" subjects represent the summation of two effects which do not reach significance alone but do so in combination, namely a greater difficulty in learning for "high-anxiety" subjects and a greater difficulty in changing from one response to another.

TABLE I

Mean Trials to reach Criterion in Part II

	"high-anxiety"	*"low-anxiety"*	*Total*
Experimental	26·0	14·3	20·15
Control	16·4	15·7	16·05
Total	21·20	15·00	18·10

Bartlett's test for homogeneity of variance was used on the data for Part II, and the assumption of homogeneity was not confirmed ($chi^2 = 17{\cdot}5$, $p<{\cdot}05$). It was therefore held advisable to supplement the analyses of variance and covariance with a non-parametric test. Wilcoxon's T-test (Wilcoxon, 1949; Moses, 1952) was chosen and applied to the results for Part II. In view of our unidirectional predictions, a one-tailed test was appropriate. The difference between "high-anxiety" and "low-anxiety" groups turned out to be significant at the ·02-level on a two-tailed test and at the ·01-level on a one-tailed test (Rank totals (T) = 327,493). The interaction initial anxiety-level x instructions (experimental or control) fell just below the ·05-level on a two-tailed test but exceeded it on a one-tailed test (T = 363,457). Our two predictions are thus once again confirmed.

IV

Discussion

While recommendations for a wide range of real-life problems can scarcely be based with any confidence on the results of this experiment alone, we have one more study to show, like numerous previous ones, that a high-anxiety level is not conducive to a change to more adaptive behaviour from behaviour that is no longer rewarded. In the situation represented by our experiment, it is clear that initially anxious individuals find difficulty in adopting new response-patterns when they are exposed to alarming exhortations. But if they are reassured, they can be as flexible as their less anxious fellows. In everyday life, people are frequently subjected to procedures intended to discomfort or frighten them as a means of inducing a radical alteration in their behaviour. And anxious, tense persons, whose personality makes them more prone than others to fixation on undesirable or inadequate responses, are especially likely to receive this treatment in family, scholastic, industrial or military environments.

References

CUTTS, J. R. (1953). The effect of the duration of the unconditioned stimulus on the learned fear drive in human subjects, with particular reference to the law of effect. Unpub. M.A. thesis, University of Aberdeen.

FARBER, I. E. (1948). Response fixation under anxiety and non-anxiety conditions. *J. exp. Psychol.*, **38**, 111-131.

FARBER, I. E. and SPENCE, K. W. (1953). Complex learning and conditioning as a function of anxiety. *J. exp. Psychol.*, **45**, 120-125.

HILGARD, E. R. (1951). Methods and procedures in the study of learning. In STEVENS, S. S. (Ed) *Handbook of Experimental Psychology*, New York.

HULL, C. L. (1943). *Principles of behaviour*. New York.

HULL, C. L. (1952). *A Behaviour System*. New Haven.

KATZ, D. (1948). *Psychological Atlas*. New York.

MAIER, N. R. F. (1949). *Frustration: The study of behaviour without a goal*. New York.

MCGEOCH, J. A. and MCKINNEY, F. (1937). Studies in retro-active inhibition: VIII. The influence of the relative order of presentation of original and interpolated paired associates. *J. exp. Psychol.*, **20**, 60–83.

MONTAGUE, E. K. (1953). The role of anxiety in serial rote learning. *J. exp. Psychol.*, **45**, 91–96.

MOSES, L. E. (1952). Non-parametric statistics for psychological research. *Psychol. Bull.*, **49**, 122–142.

MOWRER, O. H. (1950) *Learning theory and Personality Dynamics*. New York.

SPENCE, K. W. and FARBER, I.E. (1953). Conditioning and extinction as a function of anxiety. *J. exp. Psychol.*, **45**, 116–119.

SPENCE, K. W. and TAYLOR, J. A. (1951). Anxiety and strength of the UCS as determiners of the amount of eyelid conditioning. *J. exp. Psychol.*, **42**, 183–188.

TAYLOR, J. A. (1951). The relationship of anxiety to the conditioned eyelid response. *J. exp. Psychol.*, 41, 81–92.

TAYLOR, J. A. (1953). A Personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, **48**, 285–290.

TAYLOR, J. A. and SPENCE, K. W. (1952). The relationship of anxiety level to performance in serial learning. *J. exp. Psychol.*, **44**, 61–64.

WILCOXON, F. (1949). *Some rapid approximate statistical procedures*. New York.

WILCOXON, H. C. (1952). Abnormal fixation and learning. *J. exp. Psychol.*, **44**, 324–334.

BOOK REVIEWS

The Geometry of Mental Measurement. By Sir Godfrey H. Thomson. University of London Press. Pp. 60. 6s. 6d.

The title of this book may entice the intending reader by the suggestion that he will be shown how to take ruler and compasses to the contents of his head; alas, it is not so fascinating as that. What the author gives is a clear geometrical exposition of the processes involved in factor analysis by the centroid method, and for factor analysts who are already familiar with the concepts of mental testing and with the aims and uses of such analysis it should perhaps be prescribed reading. For the ordinary reader it lacks completeness in that it does not give any hint of the reasons for factorizing, either in general or in this particular way. Is it merely to get a set of independent tests, each measuring something different? If so, what is the value of explaining the results of a given number of actual tests in terms of a larger number of imaginary ones? And why should we choose common factors that explain as little as possible of the observed results, which is what is done by minimising the communalities? These are the questions he will ask himself; and as they are the questions about which most arguments among factor analysts and statisticians revolve it would be difficult to deal with them shortly. It might, however, be worth giving explicitly the basic assumption, that a subject's score on a test can be regarded as the weighted sum of his scores on a set of other tests, each measuring something different and therefore not correlated with each other. Such tests do not exist but, if the assumption be accepted, a suitable set can be deduced from the observed correlations. The size of this set will be determined to a large extent by the values of the communalities, that is, by how far we can regard the existing tests as already measuring independent effects. For the use of factors in scientific explanation the reader might be referred to Braithwaite, and, for some notion of the disputes about meaning and methods, to the discussion following Kendall's and Babington Smith's papers to the Royal Statistical Society.

The budding geometer, to whom this book is also commended in the Preface, may complain that it is incomplete for another reason. He will see no connection between the r occurring in the equation of an ellipse and the r defined as a product-moment correlation (page 9) and, in fact, this requires a knowledge of the bivariate normal distribution and of the method of estimating the constants occurring in its equation. He will surely see no reason why selection on a test should cause a plane to move at right angles to itself (page 15). It would help such a reader to be given also the model used by statisticians in dealing with multivariate data. In this, the N subjects are represented by N orthogonal axes in N dimensions. A subject's scores on the various tests are represented by distances measured from the origin along his own axis. Then all the subjects scores on one test will fix a point in the space and the line joining this point to the origin is a vector representing the test. Two such tests will define a plane through the origin and n of them will define an n-dimensional space. The cosine of the angle between two tests will equal their product-moment correlation, as can be seen by using section Five. This model will, in the end, be similar to Thomson's, but it has the advantage of giving a simple definition of r, and the result of selection, that is the removal of some of the axes, can easily be seen by simple geometry to give the results of Section Two. The main problem is then to replace these n vectors by n or more that are right angles to each other. To be familiar with such a model is valuable not only to the young geometer but also to anyone interested in current methods of multivariate analysis.

This is not a book that, on its own, is of much value to the general psychologist. Read in conjunction with some other text book on factor analysis, say the author's own earlier book *Factorial Analysis of Human Ability*, it would help to give a simple geometrical picture of the processes involved in the extraction of factors. It would be very much improved by a short bibliography.

V. R. C.

Method and Theory in Experimental Psychology. By Charles E. Osgood. London, Geoffrey Cumberlege. Oxford University Press, 1953. Pp. vii + 800. 80s.

This book though lengthy is valuable. If Professor Osgood has not had time to be brief he has found time to digest much of the recent experimental literature in American psychology, and by deliberate selection he has succeeded in presenting the individual endeavours of innumerable experimenters as a continuous story with a plot made up of a few major theoretical issues.

The division of material is traditional. There are four parts, Sensory Processes, Pp. 1–189; Perceptual Processes, Pp. 191–297; Learning, Pp. 299–599; and Symbolic Processes, Pp. 601–727. The writer's intention, as stated in his preface is to provide "a text that evaluates experimental literature in close relation to critical theoretical issues." Though there may be reservations about Osgood's selection of "critical theoretical issues" he has succeeded in his efforts to link experimental facts to theory, whilst still providing "a more detailed description of critical experiments" than is usually given. The excellent accounts and comments on experiments, supplemented by diagrams, are a feature of the book and there is no doubt that with it in his possession many a student will now contentedly leave original papers to dust on the library shelves.

On so many of the issues which he raises Osgood is eminently sound—on the psychophysical methods, his chapters on retention or on neurophysiology and learning where traditional terms such as equipotential, vicarious functioning etc. are discussed with an untraditional clarity. He is at his best where he is extending his own Statistical Theory of Figural After-Effects to cover movement perception and binocular fusion, or in his chapter on "Serial and Transfer Phenomena" relating the hoary problems of the similarity paradox and retroactive inhibition to his own work.

In such sections there is much to stimulate the thoughtful reader—it is as a contribution to Method and Theory in Experimental Psychology that the book is disappointing. Osgood writes much about theories but little about theory; his book illustrates many but it never clarifies the standpoint from which it judges one. Osgood himself favours an intermediate position—he makes the best of different schools, solving his problems by a better use of the tools which others have left lying around. It is his role to resolve the theoretical issues, say, between Gestalt and Behaviouristic psychology, and he puts forward his mediation hypothesis to provide a consistent account of adaptive behaviour which he observes has affinities with Hull, Tolman and Guthrie. It is commendable that such a reconciliation obviously inspires him, "We seem to be approaching a point where they (major learning theories) can be viewed within a consistent frame of reference" (p. 410). The student who has acquired a working knowledge of current American theory will be refreshed by Osgood's matter-of-fact comments on these issues. But in a book under this title one might have expected more appreciation of the different types of theory which do exist in present-day psychology. Osgood may be trying to reconcile the irreconcilable though he gives no hint that there are different kinds of theory. Yet under the title of theory psychologists are accustomed to find for example mere restatements of the problem in a descriptive phraseology, speculations about the operation of the nervous system, and logically inter-related models attempting to explain the workings of a system. If the different authors have not always been precise as to what they were attempting it is the task of a text which deals so exhaustively with the individual theories to be clear on this issue.

Closely related to this defect is another. Just as Osgood never hints that a different kind of theory might be possible, so too he is content to deal exclusively with the problems of psychology as defined by his contemporaries. It is not the point that a book by a single author is not all-embracing, but that significant trends are ignored. Gibson's work (1951) is ignored presumably because of the time interval between writing and publication, and perhaps the same excuse might be advanced for the uncritical acceptance of Senden's evidence and Riesen's experiments as supporting perceptual learning with no mention of the objections which largely vitiate this conclusion. But if such a book can afford to omit recent work on the study of skills it would seem ill advised to ignore the studies of comparative ethologists—here there are some traditional methods being effectively employed and a theory in an interim stage of development with important implications for the study of motivation.

However it is stressed that these criticisms are put forward only on the assumption that Osgood has written a first-class text book. It is an excellent appraisal of psychological experiments, both old and new, it has original contributions to make in several fields and is unusually sound in many others. And if it is not a classic of English prose, at least the jargon of the seminar is diluted with that of the sidewalk. As it will be widely read, another edition might take the opportunity to rectify some of the typographical errors—for example the careful reader will be unable to find any ellipses in Figure 77C, p. 214, will be puzzled by the last sentence on p. 521, will find the second reference to Figure 177 on p. 538 should be to Figure 176, and be confronted by an ordinal series of 1, 2, 4 on p. 519.

H. K.

INDEX OF AUTHORS